

Bestimmung des Fehlerverhaltens einer stereobasierten visuellen Odometrie

Bachelorarbeit

Nikolaus Fischer



BACHELORARBEIT

BESTIMMUNG DES FEHLERVERHALTENS EINER STEREOBASierten VISUELLEN ODOMETRIE

Freigabe:

Der Bearbeiter:

Unterschriften

Nikolaus Fischer



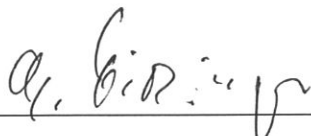
Betreuer:

Annett Stelzer



Der Institutsdirektor

Prof. Dr. G. Hirzinger



Dieser Bericht enthält 64 Seiten, 29 Abbildungen und 2 Tabellen

Fakultät für Geoinformation



Studiengang : Geoinformatik und Satellitenpositionierung
Winter-/Sommersemester : 2010/2011

B A C H E L O R A R B E I T

Bestimmung des Fehlerverhaltens einer stereobasierten visuellen Odometrie

von

Nikolaus Fischer

angefertigt am

Deutschen Zentrum für Luft- und Raumfahrt e.V.

Institut für Robotik und Mechatronik

Oberpfaffenhofen

Betreuer der Hochschule : Prof. Dr.-Ing. Peter Krzystek
Betreuer des Instituts : Dipl.-Ing. Annett Chilian

Kurzzusammenfassung

Als visuelle Odometrie bezeichnet man die Schätzung der Eigenbewegung einer Kamera aus den Bildinformationen. Aufgrund ihrer Eigenschaften wird diese Navigationsmethode oft in der mobilen Robotik angewandt. Neben den eigentlichen Messwerten ist hier auch oft die Genauigkeit der Messung erforderlich. Die Schätzung wird über markante Punkte im Bild berechnet und gibt zugleich einen geschätzten Fehler aus. Diese Arbeit untersucht den Einfluss bestimmter Eigenschaften der Punkte auf die Genauigkeit der Schätzung. Außerdem soll die Fehlerschätzung verifiziert werden. Dazu werden geeignete Experimente entwickelt und eine spezielle Testumgebung aufgebaut. Am Ende kann daraus ein Fehlerverhalten der visuellen Odometrie abgeleitet werden, welches Aufschluss über die Auswirkung der Bewegungsgröße der Stereokameras, der Verteilung der Punkte im Bild, der Anzahl der verwendeten Punkte für die Schätzung und des Abstands der Kameras zu der aufgenommenen Szene auf die Genauigkeit gibt.

Inhaltsverzeichnis

Kurzzusammenfassung.....	ii
Inhaltsverzeichnis.....	iii
Abbildungsverzeichnis.....	vi
Tabellenverzeichnis.....	viii
1 Einleitung.....	1
1.1 Motivation	2
1.2 Stereobasierte visuelle Odometrie	3
1.3 Ziel der Arbeit	3
1.4 Aufbau der Arbeit.....	4
2 Zusammenfassung der visuellen Odometrie	5
2.1 Erklärung der Formelzeichen	5
2.2 Das Stereokameramodell.....	6
2.2.1 Rektifizierung der Kamerabilder.....	6
2.2.2 Das Lochkameramodell	7
2.3 Der Stereoalgorithmus.....	9
2.3.1 Erstellen von Disparitätsbildern	9
2.3.2 Rekonstruktion der Objektkoordinaten.....	10
2.4 Der Egomotionalgorithmus.....	10
2.4.1 Das Featurematching-Verfahren.....	11
2.4.2 Die Bewegungsschätzung	11
2.5 Fehlerschätzung der visuellen Odometrie	12

3 Aufbau der Testumgebung	13
3.1 Roboterarm als Referenz	13
3.2 Installation der Stereokameras	14
3.3 Erklärung der Koordinatensysteme	15
3.4 Hand-Auge Kalibrierung	16
4 Konzeption und Durchführung der Experimente	18
4.1 Verwendete Texturen	18
4.2 Art der Bewegungen	21
4.3 Datenaufnahme	22
4.4 Erwartete Ergebnisse	22
5 Rechnerische Auswertung der Ergebnisse	24
5.1 Erklärung der Formelzeichen	24
5.2 Berechnung der Bewegungsschätzung	25
5.3 Genauigkeitsvergleich mit Referenzdaten	26
5.3.1 Transformation der KR-16 Daten	26
5.3.2 Auswertung der motion.txt Dateien	27
5.3.3 Auswertung der corner.txt Dateien	28
5.3.4 Berechnung der Genauigkeiten	28
5.4 Grafische Aufbereitung in Excel	29
6 Visualisierung und Interpretation der Genauigkeiten	31
6.1 Erreichte Genauigkeit in Abhängigkeit der Größe der Bewegung	32
6.1.1 Translatorische Bewegungen	32
6.1.2 Rotatorische Bewegungen	35
6.2 Abhängigkeit der Genauigkeit von der Anzahl der korrespondierenden Corners	37
6.2.1 Translatorische Bewegungen	37
6.2.2 Rotatorische Bewegungen	40
6.3 Abhängigkeit der Genauigkeit von der Verteilung der korrespondierenden Corners im Bild	41
6.3.1 Translatorische Bewegungen	42
6.3.2 Rotatorische Bewegungen	45

6.4 Abhängigkeit der Genauigkeit vom Abstand der Stereokameras zu der Szene	46
6.4.1 Translatorische Bewegungen	46
6.4.2 Rotatorische Bewegungen.....	48
6.5 Vergleich der tatsächlichen Fehler mit Fehlerschätzung des Egomotionalgorithmus.....	50
6.5.1 Translatorische Bewegungen	51
6.5.2 Rotatorische Bewegungen.....	52
6.5.3 Zusammenfassung der Fehlerschätzung.....	52
6.6 Zusammenfassung des Fehlerverhaltens des Egomotionalgorithmus.....	53
 7 Zusammenfassung und Ausblick	56
7.1 Zusammenfassung der Arbeit	56
7.2 Ausblick	57
 Literaturverzeichnis	59
 Anhang.....	60
A.1 Daten mit variierender texturierter Fläche.....	61
A.2 Daten mit variierender Anzahl der Corners	62
A.3 Daten mit Vergrößerung des Abstands der Stereokameras zu Szene	63
 Eigenständigkeitserklärung	64

Abbildungsverzeichnis

1.1	Der DLR-Krabbler und dessen möglicher Einsatz auf dem Mars	1
2.1	schematische Darstellung der planaren Rektifizierung.....	7
2.2	linkes Kamerabild und gerechnetes Tiefenbild	10
2.3	schematische Darstellung für das Featurematching-Verfahren in aktuellem und vorangegangenem Bild	11
3.1	Achsen des Roboterarms KR-16	14
3.2	Montierte Stereokameras auf TCP	15
3.3	Übersicht des Versuchsaufbaus und der Koordinatensysteme tcp , K und W	16
4.1	unterschiedliche Größen der texturierten Korkflächen	20
6.1	Translation in z-Richtung, Abhängigkeit der Residuen von Bewegungsgröße	32
6.2	Translation in x-Richtung, Abhängigkeit der Residuen von Bewegungsgröße	33
6.3	Rotation um z-Achse, Abhängigkeit Residuen von Bewegungsgröße.....	35
6.4	Rotation um y-Achse, Abhängigkeit Residuen von Bewegungsgröße.....	36
6.5	Translation in x-Richtung, Abhängigkeit Residuen von Corneranzahl	38
6.6	Translation in x-Richtung, Abhängigkeit Residuen von Corneranzahl mit	38
	Standardabweichung der visuellen Odometrie zu Mittelwert	38
6.7	Translation in z-Richtung, Abhängigkeit Residuen von Corneranzahl, sortiert nach Richtung und Bewegungsgröße.....	39
6.8	Rotation um z-Achse, Abhängigkeit Residuen von Corneranzahl	40
6.9	Rotation um z-Achse, Abhängigkeit Residuen von Corneranzahl, sortiert nach Richtung und Bewegungsgröße	41
6.10	Translation in x-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche	42

6.11 Translation in z-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche, sortiert nach Richtung und Bewegungsgröße	43
6.12 Translation in x-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche, sortiert nach Richtung und Bewegungsgröße	44
6.13 Rotation um z-Achse, Abhängigkeit Residuen von Größe der texturierten Fläche	46
6.14 Translation in z-Richtung, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zur Szene, sortiert nach Richtung und Bewegungsgröße	47
6.15 Translation in x-Richtung, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zur Szene, sortiert nach Richtung und Bewegungsgröße	48
6.16 Rotation um z-Achse, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zu Szene, sortiert nach Richtung und Bewegungsgröße	49
6.17 Rotation um y-Achse, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zu Szene, sortiert nach Richtung und Bewegungsgröße	49
6.18 Vergleich der Zunahme der Dreidimensionalität bei Abstandsvergrößerung	50
6.19 Translation in z- und x-Richtung, Vergleich von realen und geschätzten Residuen der visuellen Odometrie, sortiert nach Richtung und Bewegungsgröße	51
6.20 Rotation um z- und y-Achse, Vergleich von realen und geschätzten Residuen der visuellen Odometrie, sortiert nach Richtung und Bewegungsgröße	53

Tabellenverzeichnis

4.1	Vergleich von verschiedenen Texturen	19
4.2	Fläche der unterschiedlichen Korktexturen im digitalen Bild	21

Kapitel 1

Einleitung

Die mobile Robotik gewinnt in unserer Zeit immer mehr an Bedeutung. Durch die Automatisierung in vielen Bereichen der Industrie und des privaten Lebens werden immer mehr Systeme benötigt, die selbstständig und zuverlässig arbeiten können. Der Bereich der Robotik ist eng verbunden mit der Navigation. Ein mobiler Roboter sollte sich ohne Hilfe im Raum orientieren können, dabei seine Umgebung erkennen, und diese Informationen interpretieren, um weitere Entscheidungen zu treffen. Um die Umgebung zu erkennen werden häufig einzelne Kameras oder Stereokameras verwendet. Die visuelle Odometrie ist eine Möglichkeit sich mit Stereokameras durch den Raum zu navigieren. Dabei wird eine Relativbewegung der Kameras zwischen zwei aufeinanderfolgenden Bildern berechnet. Diese Technik wird nicht nur auf der Erde eingesetzt, sondern z.B. auch für die Exploration von fremden Planeten verwendet. Das bekannteste Projekt in diesem Bereich war die Mission der NASA-Rover Spirit und Opportunity. Sie waren in der Lage sich autonom auf dem Mars zu bewegen und dort ihre Umgebung zu erkunden.

(a) Krabbler in Testumgebung



(b) Vision des Explorationsroboters



Abbildung 1.1 Der DLR-Krabbler und dessen möglicher Einsatz auf dem Mars, Quelle: DLR

Am Institut für Robotik und Mechatronik am Deutschen Zentrum für Luft- und Raumfahrt in Oberpfaffenhofen (DLR) wird zur Zeit an einem mobilen sechsbeinigen Laufroboter geforscht. Der „Krabbler“ soll in Zukunft als „Search and Rescue“-Roboter in Katastrophengebieten, oder als Explorationsroboter auf anderen Planeten, wie zum

Beispiel dem Mars, eingesetzt werden (siehe Abbildung 1.1). Mit Hilfe von Stereokameras und der visuellen Odometrie soll er in der Lage sein, sich autonom durch unebenes Gelände zu navigieren und seine Umgebung zu kartieren [1].

Für solche Verfahren benötigt der Roboter nicht nur die aufgenommenen Messwerte, sondern auch eine Aussage über die Zuverlässigkeit jeder einzelnen Messung. Die vorliegende Arbeit soll genauere Informationen über das Fehlerverhalten der visuellen Odometrie liefern, um dem Roboter die Möglichkeit zu geben, seinen eigenen Positions- und Orientierungsfehler abzuschätzen.

1.1 Motivation

Zur Schätzung seiner Position und Orientierung besitzt ein Roboter in der Regel mehrere verschiedene Sensoren. Der Krabblер verwendet zusätzlich zu den Stereokameras Gelenksensoren in den Beinen, über die eine Beinodometrie berechnet werden kann, sowie eine inertielle Messeinheit (IMU). Um robuste Lageinformationen zu erhalten, werden die Messwerte möglichst aller Sensoren zu einer genauen Positions- und Orientierungsschätzung fusioniert. Allerdings zeigen die Sensoren unterschiedliche Genauigkeiten der Werte auf, die oft durch die Umgebung beeinflusst werden. Für eine gute Fusion der Messwerte ist es daher wichtig, nicht nur die Sensormessungen selbst zu kennen, sondern auch deren Genauigkeit. Damit kann ein höheres Gewicht auf genaue Messungen gelegt werden und ungenauere Messwerte haben einen geringeren Einfluss auf das Fusionsergebnis.

Des Weiteren ist es für fast alle Aufgaben in der Robotik wichtig, den Fehler einer Messung so gut zu kennen, wie die Messung selbst. Auf dieser Grundlage können richtige Entscheidungen getroffen werden. Nähert sich der Roboter zum Beispiel einem Objekt, welches er zuvor kartiert hat, kann er über die Genauigkeitsinformation seiner Position und Orientierung entscheiden, mit welchem Abstand er das Objekt umfährt.

Daher ist es relevant, das Fehlerverhalten der visuellen Odometrie genau zu erforschen. Ein großer Einfluss auf die Genauigkeit ist die aufgenommene Szene der Stereokameras. Markante Punkte in den aufgenommenen Bildern, sogenannte Features, sind Grundlage für die Berechnung der visuellen Odometrie. Der Hauptteil dieser Arbeit soll zeigen, wie sich bestimmte Eigenschaften der Features auf die Genauigkeit der berechneten Position auswirkt. Dieses Feld wurde in Bezug auf die visuelle Odometrie bisher kaum untersucht. Das Institut für Robotik und Mechatronik bietet die Möglichkeit, eine Testumgebung anzulegen, in der das Fehlerverhalten

ausführlich untersucht werden kann. Dabei wird der hochgenaue Roboterarm KR-16 von KUKA als Referenzsystem benutzt. Dieser kann mit Stereokameras verbunden werden. Für die Berechnung der robusten visuellen Odometrie kann die Stereolibrary von Dr. Heiko Hirschmüller verwendet werden. Diese wird bereits auf vielen unterschiedlichen robotischen Systemen eingesetzt.

1.2 Stereobasierte visuelle Odometrie

Als visuelle Odometrie bezeichnet man die Schätzung der Eigenbewegung einer Kamera aus den Bildinformationen. Stereokamerabasierte Verfahren bieten den Vorteil, dass die 3D Koordinaten der Bildpunkte rekonstruiert werden können. Das ist mit Hilfe eines Disparitätsbilds möglich. Es entsteht aus dem Vergleich der Pixelwerte von linkem und rechtem Kamerabild und liefert ein Tiefenbild des linken Kamerabilds. Durch das Matching von Featurepunkten in aufeinanderfolgenden Kamerabildern erhält man zwei korrespondierende Punktwolken. Die Transformation, welche die Punktwolken ineinander überführt, ist die Eigenbewegung der Kamera. Durch das Aufsummieren der Relativbewegungen entsteht eine Absolutbewegung, ausgehend von dem Startpunkt der Stereokameras. Da es sich hierbei um ein inkrementelles Verfahren handelt, wird dafür auch der Begriff visuelle Odometrie verwendet.

Die visuelle Odometrie hat den Vorteil, dass es ein passives System ist, und keine aktiven Signale aussenden muss, um die Umgebung zu erfassen. Daher eignet sich ein Einsatz in Gebieten, wo die klassische Navigation z.B. mit GPS nicht mehr möglich ist. Vor allem die Robotik ist daher eng gekoppelt mit der visuellen Odometrie. Neben der Navigation können aus den Bildern mit Hilfe der Bildverarbeitung weitere Informationen aus der Umgebung gewonnen werden. Durch das inkrementelle Verfahren der visuellen Odometrie summieren sich allerdings alle Fehler in der absoluten Bewegungsschätzung auf. Deswegen wird diese oft mit weiteren Sensoren zur Navigation verbunden, wie z.B. Neigungssensor oder einem Kompass.

1.3 Ziel der Arbeit

Das Ziel der Arbeit ist, eine Aussage über die Genauigkeit der visuellen Odometrie von Dr. Heiko Hirschmüller in Zusammenhang mit den für die Berechnung verwendeten Features zu treffen. Außerdem liefert der Algorithmus selbst auch eine

Fehlerschätzung der durchgeführten Relativbewegung. Die Fehlerschätzung wurde parallel zu dieser Arbeit implementiert und ist daher noch nicht experimentell verifiziert worden. Sie soll daher bei der Genauigkeitsuntersuchung mit überprüft werden.

Dafür wird zuerst eine geeignete Testumgebung, bestehend aus den Stereokameras und dem Roboterarm aufgebaut. Damit sollen Experimente entwickelt werden, die die Genauigkeit der Bewegungsschätzung untersuchen können. Dabei soll die Auswirkung der Anzahl und der Verteilung der korrespondierenden Features auf die Genauigkeit untersucht werden. Außerdem soll der Einfluss des Abstands der Kameras zu der aufgenommenen Szene auf das Fehlverhalten aufgezeigt und der geschätzte Fehler der visuellen Odometrie mit dem tatsächlichen Fehler verglichen werden. Die Erkenntnisse aus den Experimenten sollen am Ende der Arbeit ein Fehlverhalten für die visuelle Odometrie des DLR beschreiben.

1.4 Aufbau der Arbeit

Das 2. Kapitel gibt einen Überblick über die Funktionsweise des Stereo- und Egomotionalgorithmus von Dr. Heiko Hirschmüller, welche die Grundlage für diese Arbeit sind. Das 3. Kapitel geht auf den Aufbau der Testumgebung, die für diese Arbeit entworfen wurde, ein. Hier werden die technischen Daten des Roboterarms KR-16, sowie die Daten der Stereokameras und deren Verbindung mit dem Arm erläutert. Darüber hinaus wird die spezielle Hand-Auge Kalibrierung der beiden verbundenen Systeme beschrieben. Kapitel 4 geht auf die Konzipierung der Experimente ein, die in der Testumgebung durchgeführt werden. Damit die Experimente eine Abhängigkeit der Fehler von den Eigenschaften der detektierten Features aufzeigen können, müssen verschiedene Faktoren berücksichtigt werden. Hier werden die Auswahl einer geeigneten texturierten Fläche, auf die die Stereokameras in den Experimenten gerichtet sind, sowie die Bewegungsarten erläutert. Außerdem werden die Datenaufnahmen und die erwarteten Ergebnisse der Experimente beschrieben. Kapitel 5 erklärt die Aufbereitung der Daten aus den Experimenten, damit sie in Kapitel 6 interpretiert werden können, und ein Fehlverhalten der visuellen Odometrie daraus abgeleitet werden kann. Das letzte Kapitel fasst die Ergebnisse der Arbeit zusammen und gibt einen Ausblick für zukünftige Experimente.

Kapitel 2

Zusammenfassung der visuellen Odometrie

Grundlage für diese Arbeit ist die auf C basierende Stereolibrary von Dr. Heiko Hirschmüller [2]. Die Library besteht aus einem Stereoalgorithmus, der ein Tiefenbild der aufgenommenen Szene berechnet und einem Egomotionalgorithmus, der die relative Bewegung zwischen zwei Bildern schätzt. Über die Aufsummierung der relativen Translationen und Rotationen entsteht eine visuelle Odometrie, die die Bewegung des Systems im dreidimensionalen Raum, ausgehend von der Startposition der Stereokameras, berechnet.

Abschnitt 2.1 erklärt die Formelzeichen, welche in diesem Kapitel verwendet werden. Das zugrunde liegende Kameramodell der Algorithmen wird in Abschnitt 2.2 erläutert. Abschnitt 2.3 erklärt, wie aus den Stereobildern ein Disparitätsbild gerechnet wird. Wie anhand des Modells und den aufgenommenen Stereobildern die visuelle Odometrie geschätzt wird, behandelt Abschnitt 2.4. Die Funktionsweise der Fehlerschätzung der berechneten Bewegungsschätzung beschreibt Abschnitt 2.5.

2.1 Erklärung der Formelzeichen

Allgemeine Formelzeichen:

O	Ursprung eines Koordinatensystems
T	allgemeiner Translationsvektor
R	allgemeine Rotationsmatrix
C, P	Punkte im Kamerakoordinatensystem
Q	Punkt im Weltkoordinatensystem
p	Punkt im Bildkoordinatensystem
p'	Punkt im rektifizierten Bild
b	Bildbreite in Pixeln
h	Bildhöhe in Pixeln
f	Brennweite in Pixeln
A	Projektionsmatrix

s	Skalierungsfaktor
t	Verschiebungsvektor
ε_{gi}	geometrischer Fehler

Indizes:

r	rechts
l	links
S	Stereokamerasystem
d	Disparität
x	x-Wert
y	y-Wert
rel	relativ

2.2 Das Stereokameramodell

Das Stereokameramodell wird benötigt, um einen Punkt vom Welt - in das Bildkoordinatensystem zu transformieren [2, S.6-20]. Ein reales Kameramodell besteht aus einem linearen Teil, der für die Transformation der Punkte verantwortlich ist, und einem nichtlinearen Teil, der die Verzeichnungen durch die Linse der Kamera korrigiert [2, S.6]. Dieses Modell wird in der Doktorarbeit von Dr. Heiko Hirschmüller ausführlich beschrieben. Da das Modell sehr komplex ist und aus 20 Parametern besteht, wurde es über die planare Rektifizierung für den Algorithmus zu einem idealen Lochkameramodell vereinfacht, um eine bessere Performance zu bekommen.

2.2.1 Rektifizierung der Kamerabilder

Durch die planare Rektifizierung der Bilder kann ein Lochkameramodell verwendet werden, das von einer perfekten Kamera ohne Linsenverzeichnung ausgeht. Dabei werden linkes und rechtes Kamerabild, die zueinander verdreht sind, auf eine gemeinsame Bildebene projiziert. Dazu wird die Linie des Lichtstrahls des Punktes im Weltkoordinatensystems durch das optische Zentrum der Kameras (o_l oder o_r) simuliert. Der Schnittpunkt dieser Linie mit der gemeinsamen Ebene ergibt den Punkt

p' im rektifizierten Bild, wie Abbildung 2.1 zeigt. Die Epipolarlinie ist die Linie, wo sich die Bildfläche mit der Ebene, die durch Objektpunkt und die beiden Projektionszentren der Kameras aufgespannt wird, schneidet. Ziel der planaren Rektifizierung ist, dass die Epipolarlinien in beiden Bildern parallel werden. Das Verfahren der Rektifizierung wird auf jedes aufgenommene Bild angewandt.

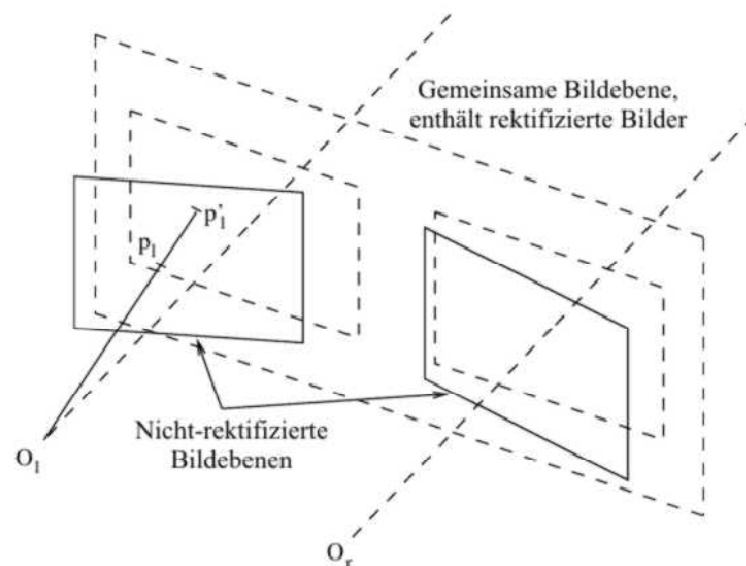


Abbildung 2.1 schematische Darstellung der planaren Rektifizierung, in [2, S. 13]

Die Parameter für die Rektifizierung werden über eine Kalibrierung der Stereokameras ermittelt. Dabei wird eine intrinsische Kalibrierung durchgeführt, in der jede Kamera für sich kalibriert wird, um die Verzeichnungen zu ermitteln. Bei der extrinsischen Kalibrierung werden die beiden Kameras gleichzeitig kalibriert, um die Transformation zwischen den Einzelkameras zu bestimmen. Für diese Arbeit wurden alle Kalibrierungen gemeinsam mit der Hand-Auge Kalibrierung durchgeführt, die in Abschnitt 3.4 noch weiter erläutert wird.

2.2.2 Das Lochkameramodell

Formel (2.1) beschreibt die lineare Transformation eines Punktes Q im Weltkoordinatensystem mit Hilfe einer Rotationsmatrix R und einem Translationsvektor T in einen Punkt P im Kamerakoordinatensystem. Diese wird auch als extrinsische Transformation bezeichnet.

$$P = R \cdot Q + T \quad (2.1)$$

Daraus ergeben sich sechs extrinsische Parameter, die sich aus drei Rotationen und drei Translationen zusammen setzen.

Formel (2.2) zeigt die extrinsische Transformation zwischen den Stereokameras und stellt eine Beziehung zwischen den beiden Kameras her. Sie beschreibt die Transformation eines Punktes im linken Kamerasystem P_l in das rechte System P_r . Die Werte von T_s und R_s ändern sich nicht mit der Bewegung der Kameras.

$$P_r = R_s \cdot (P_l - T_s) \quad (2.2)$$

$$\text{wobei } R_s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ und } T_s = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix}$$

Der Basisabstand t steht für die Verschiebung der Kameras entlang der x-Achse. Um einen Punkt des Weltkoordinatensystems letztendlich im Bildkoordinatensystem richtig darzustellen, müssen die Punkte des linken und rechten Kamerasystems mit Hilfe der intrinsischen Transformation überführt werden. Mit den rektifizierten Bildern kann der Stereoalgorithmus für eine Bewegungsschätzung davon ausgehen, dass die Kameras ideale Lochkameras ohne Verzeichnungen sind. Dadurch sind nur noch zwei intrinsische Parameter notwendig, die Brennweite f , die für beide Kameras gleich ist, und der Bildmittelpunkt $(\frac{b}{2}$ und $\frac{h}{2})$. b ist dabei die Bildbreite und h die Bildhöhe. Aus diesen Parametern ergibt sich die Projektionsmatrix

$$A = \begin{pmatrix} f & 0 & \frac{b}{2} \\ 0 & f & \frac{h}{2} \\ 0 & 0 & 1 \end{pmatrix} \quad (2.3)$$

Wie aus einem verzeichnungsfreien Kamerapunkt P der linke Bildpunkt p_l und der rechte p_r entsteht, zeigt die Formel (2.4) .

$$s_1 \cdot p_l = A \cdot P, \quad s_2 \cdot p_r = A \cdot R_s \cdot (P - T_s) = A \cdot (P - T_s) \quad (2.4)$$

Anhand der Skalierungsfaktoren s_1 und s_2 werden die z-Koordinaten von p immer auf 1 gesetzt, da das Bildkoordinatensystem nur zweidimensional ist.

Die daraus resultierenden Eigenschaften sind für die Stereovision sehr von Nutzen: Die Abbildung eines Punktes im Weltkoordinatensystem muss im linken und im rechten Stereobild in der gleichen Bildzeile sein. Somit kann die spätere Suche nach

übereinstimmenden Pixeln in Abschnitt 2.2.1 im eindimensionalen Raum durchgeführt werden. Des Weiteren muss die Bildspalte des Punktes im rechten Bild niedriger oder gleich sein, wie die im linken Bild. Durch diese Eigenschaft ist es möglich, Disparitätsbilder effizient zu erzeugen.

2.3 Der Stereoalgorithmus

Der Stereoalgorithmus errechnet ein Tiefenbild aus den Bildern der linken und rechten Kamera. Anhand dieser Daten kann ein Disparitätsbild gerechnet werden, das Grundlage für die Rekonstruktion der Objektkoordinaten ist.

2.3.1 Erstellen von Disparitätsbildern

Um ein Disparitätsbild erstellen zu können, wird zuerst ein Stereomatching durchgeführt. Das bedeutet, dass für jedes Pixel im linken Kamerabild das zugehörige im rechten Bild mit Hilfe einer Korrelationsfunktion gefunden werden muss. Durch die Eigenschaften der rektifizierten Bilder muss für das Matching nur in der gleichen Bildzeile gesucht werden. Das schließt schon mal einen Großteil des Bildes aus, verhindert falsche Korrespondenzen und erhöht die Geschwindigkeit. Das Pixelpaar mit der maximalen Korrelation wird als korrespondierendes Paar detektiert. Der Disparitätswert p_d ergibt sich über

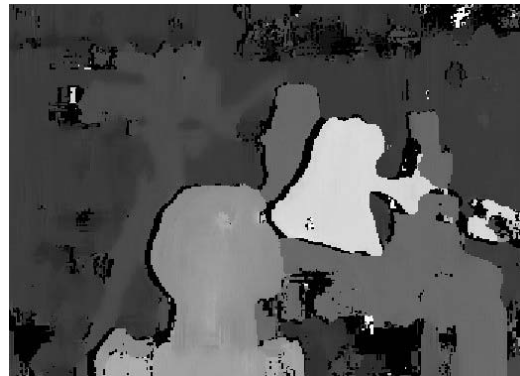
$$p_d = p_{lx} - p_{rx} \quad (2.5)$$

Die x-Werte der korrespondierenden Pixelpunkte im linken Bild p_{lx} und im Rechten p_{rx} werden subtrahiert. Somit kann für jedes Pixel im linken Kamerabild ein Disparitätswert errechnet werden. Abbildung 2.2 zeigt die Werte, die in einem Tiefenbild gespeichert werden. Je größer der Abstand der Kameras zu dem Objekt wird, desto kleiner wird der Disparitätswert und desto dunkler der Pixel im Disparitätsbild.

(a) linkes Kamerabild



(b) Tiefenbild

**Abbildung 2.2** linkes Kamerabild und gerechnetes Tiefenbild, in [2, S. 26 und S.38]

2.3.2 Rekonstruktion der Objektkoordinaten

Der dreidimensionale Punkt P im Kamerakoordinatensystem kann aus dem zweidimensionalen Bildpunkt p rekonstruiert werden [2, S.46-60]. Dazu werden die Gleichungen für das verwendete Lochkameramodell in Abschnitt 2.2.2 umgeformt und der Disparitätswert p_d aus Abschnitt 2.3.1 mit einbezogen.

$$P_x = p_x \cdot \frac{t}{p_d} \quad (2.6a)$$

$$P_y = p_y \cdot \frac{t}{p_d} \quad (2.6b)$$

$$P_z = f \cdot \frac{t}{p_d} \quad (2.6c)$$

$$\text{mit } p_x = p_{lx}, p_y = p_{ly} = p_{ry}$$

Dabei steht f für die Brennweite und t für den Basisabstand der Stereokameras.

2.4 Der Egomotionalgorithmus

Der Egomotionalgorithmus schätzt mit Hilfe von markanten Punkten und den rekonstruierten Objektkoordinaten eine Bewegung zwischen zwei aufeinanderfolgenden Bildern [2, S.61-82].

2.4.1 Das Featurematching-Verfahren

Über den Harris Eckendetektor werden markante Punkte im Bild aus starken lokalen Helligkeitsunterschieden erkannt [3]. Diese Punkte werden als „Corner“ oder „Feature“ bezeichnet. Um die Corners wird ein rechteckiges Fenster gelegt. Das Feature des aktuellen Bildes wird mit den detektierten Features des Vorgängerbildes über Korrelation verglichen (Abbildung 2.3). Das Featurepaar mit der höchsten Korrelation wird als korrespondierendes Featurepaar erkannt.

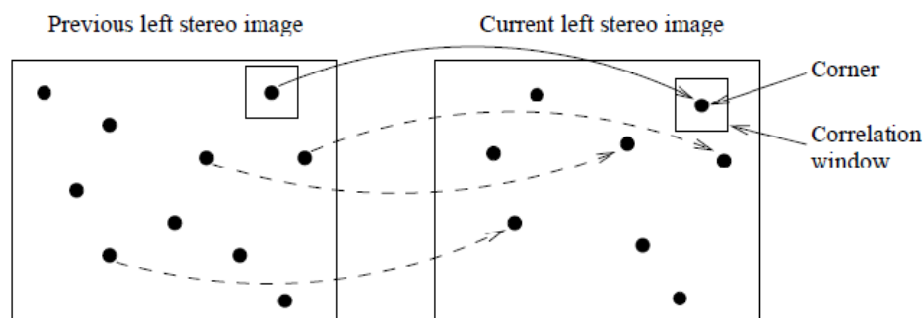


Abbildung 2.3 schematische Darstellung für das Featurematching-Verfahren in aktuellem und vorangegangenen Bild, in [2, S. 65]

2.4.2 Die Bewegungsschätzung

Anhand der korrespondierenden Punkte von einem Kamerabild auf das nächste, die über das Featurematching-Verfahren detektiert wurden, kann eine Bewegung der Stereokameras zwischen den Bildern bestimmt werden. Voraussetzung dafür ist, dass die Umgebung der Kameras statisch ist. Nach den Formeln (2.5) und (2.6) sind die dreidimensionalen Punkte aller korrespondierenden Corners bekannt.

Um möglichst früh falsche Korrespondenzen auszuschließen, werden die Abstände aller Punktpaare des aktuellen Bildes C_i und C_j , sowie die Abstände der Punktpaare des Vorgängerbildes P_i und P_j , relativ zueinander berechnet. Da die Umgebung statisch ist, dürfen sich die Relativabstände der korrespondierenden Punkte im Bild jeweils zueinander nicht ändern. Stimmen bei einem Vergleich die Abstände innerhalb einer bestimmten Toleranzgrenze überein, werden die Punkte für die Bewegungsschätzung verwendet.

Ein weiteres Verfahren, um Ausreißer bei den Korrespondenzen zu eliminieren, ist die Bedingung, dass sich die Kameras zwischen zwei Bildern nicht mehr als 45° drehen dürfen. In diesem Fall gibt es selten Überlappungen und dementsprechend korrespondierende Features in den Bildern. Wäre für ein korrespondierendes

Featurepaar eine solche Drehung notwendig, dann wird diese Korrespondenz als falsch angenommen und aus der Liste der Korrespondenzen entfernt.

Mit den verbleibenden korrespondierenden Punkten wird die Bewegungsschätzung berechnet. Sie besteht aus einer relativen Rotation und einer relativen Translation, um den Punkt des aktuellen Bildes C_i in den Punkt im Vorgängerbild P_i zu transformieren (Formel (2.7)). Dabei wird ein geometrischer Fehler ε_{gi} angenommen. Die geometrischen Fehler aller relativen Transformationen zwischen den korrespondierenden Featurepaaren werden über eine nichtlineare Optimierung minimiert. Daraus ergeben sich die relativen Transformationen R_{rel} und T_{rel} .

$$P_i = R_{rel} \cdot C_i + T_{rel} + \varepsilon_{gi} \quad (2.7)$$

Durch Aufsummierung der einzelnen Relativbewegung errechnet sich eine Absolutbewegung der Stereokameras, deren Anfangspunkt die Position der Kameras vom ersten aufgenommenen Bild ist. Da dieses Verfahren inkrementell ist, kann die Schätzung auch als visuelle Odometrie bezeichnet werden.

2.5 Fehlerschätzung der visuellen Odometrie

Die Egomotion schätzt zusätzlich zu der Bewegungsschätzung auch den Fehler der visuellen Odometrie. Über die Umkehrung der Funktion, die bei der nichtlinearen Optimierung zur Berechnung der relativen Transformation zwischen zwei aufeinanderfolgenden Bildern minimiert wird, wird berechnet, wie sich ein Bildfehler von 0,5 Pixeln bei der Featuredetektion aller Features auf die geschätzte Bewegung auswirkt. Die Funktionsweise der Fehlerschätzung wird in dem Paper [4, B.2] ausführlich beschrieben. Sie gibt einen maximalen Fehler für die relative Translation und Rotation zwischen den Bildern aus. Tatsächlich liegt die durchschnittliche Genauigkeit der Schätzung deutlich höher, da über die hohe Anzahl von korrespondierenden Cornern eine Genauigkeit im Subpixelbereich auch unter 0,5 Pixeln erreicht werden kann. Außerdem wird bei der Schätzung angenommen, dass es keine falschen Korrespondenzen bei dem Matching gibt. In der Realität gibt es aber fehlerhafte Korrespondenzen, die die Bewegungsschätzung verfälschen. Die Schätzung der translatorischen Fehler wird in mm ausgegeben. Die der rotatorischen Fehler sind als Angle-Axis-Parameterfehler angegeben. Aus diesen Parametern lässt sich ein rotatorischer Fehler in Eulerwinkeln nur im Zusammenhang mit der geschätzten Rotation berechnen.

Kapitel 3

Aufbau der Testumgebung

Um die Genauigkeit der berechneten Bewegung der visuellen Odometrie in Abhängigkeit zu den detektierten Features zu beurteilen und mit einer absoluten Referenz zu vergleichen, wird eine Testumgebung aufgebaut, die speziell auf die Problemstellung dieser Arbeit ausgelegt ist. Für das Referenzsystem wurde der am DLR verfügbare, hoch genaue Roboterarm KR-16 der Firma KUKA ausgewählt.

Abschnitt 3.1 erläutert den Aufbau und die Funktionsweise des Arms. Die für die visuelle Odometrie benötigten Stereokameras werden auf den Toolcenterpoint (TCP) am Ende des Arms montiert. Auf die technischen Details der Kameras geht Abschnitt 3.2 ein. Die verschiedenen Koordinatensysteme der Testumgebung werden in Abschnitt 3.3 zusammengefasst und erklärt. Da sich die Koordinaten des Roboters auf den TCP beziehen, müssen über eine Hand-Auge Kalibrierung des Systems die Transformationsparameter zwischen Roboter und Stereokameras bestimmt werden. Die Kalibrierung wird in Abschnitt 3.4 beschrieben. Somit können die Koordinaten bzw. die durchgeführten Bewegungen des Roboters später in das Bezugssystem der Stereokameras transformiert und direkt mit der Bewegungsschätzung der visuellen Odometrie verglichen werden.

3.1 Roboterarm als Referenz

Der Roboterarm KR-16 von KUKA ist ein elektronisch gesteuerter Industrieroboter [5]. Er besitzt sechs einzeln ansteuerbare Achsen, wie Abbildung 3.1 zeigt. Die maximale Reichweite des Arms beträgt 1,60 m. Die Traglast beläuft sich auf 16 kg. Der Roboterarm ist über die Basis fest mit dem Boden verankert. Der Kopf des Roboterarms wird als Toolcenterpoint (TCP) bezeichnet. Dort können weitere Geräte an einer fixen Vorrichtung angebracht werden. In der Testumgebung sind hier die Stereokameras mit einer speziellen Halterung, die in Abschnitt 3.2 erklärt wird, montiert. Manuelle Bewegungen des Roboterarms können über das Programmierhandgerät durchgeführt werden, automatische Bewegungen werden mit Hilfe des KUKA Software-Programms und einer systemeigenen Programmiersprache einprogrammiert [6].



Abbildung 3.1 Achsen des Roboterarms KR-16, aus [6, S.54]

Über Rotationen um die sechs Achsen sind translatorische und rotatorische Bewegungen des TCP im dreidimensionalen Raum möglich. Dabei können Punkte im Roboterweltkoordinatensystem W definiert werden, welche der Arm bei Ablauf des Programms selbstständig anfährt. Die Positionswiederholgenauigkeit des KR-16 liegt bei $\pm 0,05\text{mm}$ [5, S. 8].

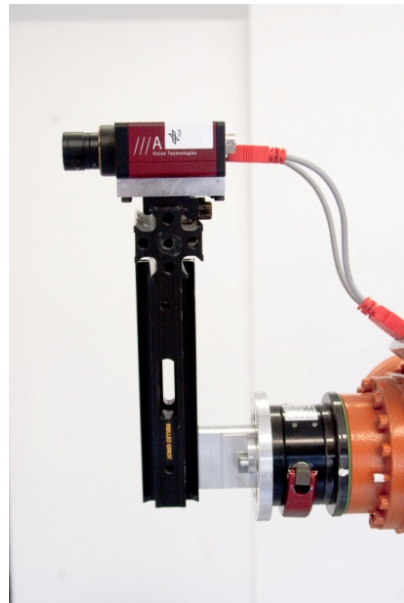
3.2 Installation der Stereokameras

Die für die visuelle Odometrie verwendeten Stereokameras stammen von der Firma Allied Vision Technologies. Die Kompaktkameras vom Typ Marlin F-201C besitzen einen zwei Megapixel CCD Sensor von Sony. Die maximale Framerate bei voller Auflösung von 1628×1236 Pixeln liegt bei 12,5 fps. Aufgrund der Datenübertragung via Firewire Kabel werden die Datenmengen bei voller Auflösung zu groß für eine problemlose Verarbeitung. Deshalb werden die Daten während der Tests für diese Arbeit mit 618×814 Pixeln und 5 fps aufgezeichnet. Die Objektive der Firma Sony vom Typ VCL-O6S12XM besitzen eine Brennweite von 6mm und eine Lichtstärke von 1:1,4. Der horizontale Öffnungswinkel liegt bei 50° , der Vertikale bei 40° . Abschnitt 3.4 beschreibt die Kalibrierung der Stereokameras. Abbildung 3.2 zeigt die Einbindung der Stereokameras in die Testumgebung. Sie werden über einen Flansch statisch mit dem TCP des Roboterarms verbunden. Während der Experimente kann sich die Verbindung nicht lösen oder verändern.

(a) frontal



(b) seitlich

**Abbildung 3.2** Montierte Stereokameras auf TCP

3.3 Erklärung der Koordinatensysteme

Die Testumgebung besitzt durch die Kombination der verschiedenen Instrumente unterschiedliche Koordinatensysteme, auf die sich gelieferte Koordinaten beziehen. Abbildung 3.3 zeigt den Versuchsaufbau schematisch mit den entsprechenden Koordinatensystemen.

Die Software des KR-16 gibt die Referenzbewegungen des Roboters als Translation und Rotation von Welt- in TCP-Koordinatensystem aus. Das kartesische Weltkoordinatensystem W ist durch den Ursprung am Roboterfuß in der Mitte der Basis des KR-16 definiert [6, S.53]. Dabei zeigt die x-Achse nach vorne, die y-Achse nach links und die z-Achse nach oben. Der Ursprung des TCP-Koordinatensystems tcp liegt im Mittelpunkt des Toolcenterpoints. Die x-Achse zeigt nach unten, die y-Achse nach links und die z-Achse nach vorne.

Die zu überprüfenden Bewegungen des Egomotion Algorithmus liegen als relative Translation und Rotation von einem Bild auf das Nächste vor. Diese beziehen sich auf das Kamerakoordinatensystem K . Der Ursprung liegt im Bildhauptpunkt der linken Stereokamera. Das ist der Punkt, in dem die optische Achse die Bildebene schneidet. Die x-Achse ist nach rechts gerichtet, die y-Achse nach unten und die z-Achse zeigt in Blickrichtung der Kameras nach vorne.

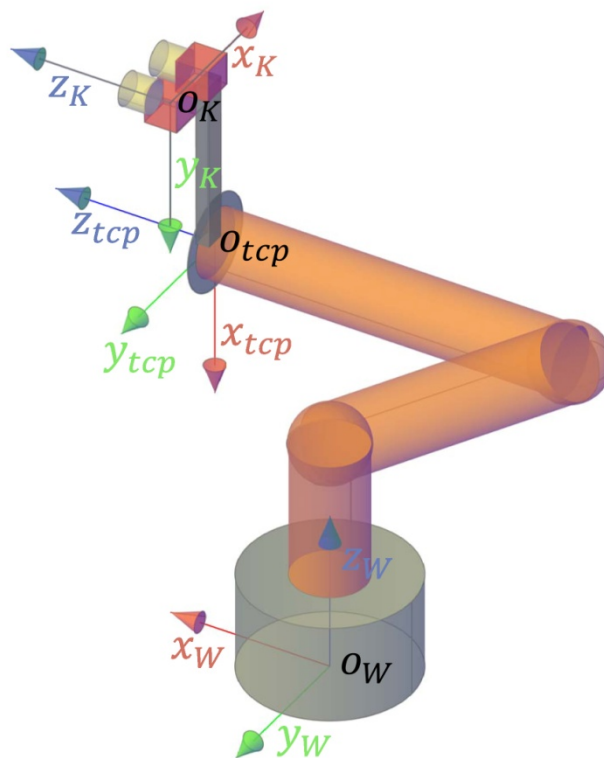


Abbildung 3.3 Übersicht des Versuchsaufbaus und der Koordinatensysteme tcp , K und W

3.4 Hand-Auge Kalibrierung

Der Versuchsaufbau wird nach dem Hand-Auge Kalibrierungsverfahren von Klaus Strobl kalibriert [7]. Durch diese spezielle Kalibrierung können Translation und Rotation zwischen dem Ursprung des KR-16 TCP- und des Kamerakordinatensystems der visuellen Odometrie bestimmt werden. Anhand der gelieferten Parameter werden bei der Datenauswertung in Kapitel 5 die Bewegungen des TCP in das Kamerakordinatensystem transformiert und mit den Berechnungen des Egomotionalgorithmus verglichen.

Für die Kalibrierung werden sechs Posen über einer Kalibrierplatte mit einem quadratischen schwarz-weiß Raster durch den Roboterarm und den montierten Stereokameras angefahren. Die Koordinaten des Roboters und die aufgenommenen Bilder der Stereokameras werden von DLR internen Programmen aufgezeichnet und gespeichert. Die Programme „calde“ und „callab“ [8,9] basieren auf dem Kalibrierungsverfahren von Klaus Strobl. Die Bilder sowie die Abmessungen der Kalibrierplatte werden zuerst in das Programm „calde“ geladen. Damit werden in den

Bildern die Eckpunkte der Quadrate der Kalibrierplatte detektiert. Deren Relativpositionen sind genau bekannt und dienen als Referenzpunkte. Anschließend werden die detektierten Punkte, sowie die Posen des Roboters in „callab“ geladen. Dieses Programm berechnet die intrinsische und extrinsische Kamerakalibrierung, sowie die Hand-Auge-Kalibrierung.

Als Ergebnis werden für die Hand-Auge-Kalibrierung der Translationsvektor $T_{K tcp}^K$ und die Rotationsmatrix $R_{tcp K}$ zwischen TCP- und Kamerakoordinatensystem ausgegeben. Diese werden zusammen mit den intrinsischen und extrinsischen Kalibrierdaten für die Rektifizierung in einer Parameterdatei gespeichert. Hier sind außerdem auch die Parameter für das Stereomatching und die Egomotionberechnung gespeichert. Nach der Kalibrierung betragen die Standardabweichungen am TCP $0,0699411^\circ$ und $0,493111\text{mm}$. Bei einer Umrechnung der Position von Kamera auf TCP über die Hand-Auge Kalibrierung entsteht demnach in 68% der Fälle ein absoluter Positionsfehler kleiner als $0,07^\circ$ und $0,49\text{mm}$.

Kapitel 4

Konzeption und Durchführung der Experimente

Ausgehend von der in Kapitel 3 beschriebenen Testumgebung wurden für diese Arbeit Experimente entwickelt, die die Bewegungsschätzung des Egomotionalgorithmus mit Hilfe der hochgenauen Positionsmessungen des KR-16 verifizieren können. Die Genauigkeit der Bewegungsschätzung ist unter anderem abhängig von der Beschaffenheit der Textur, auf die die Stereokameras gerichtet sind und mit deren Hilfe der Algorithmus korrespondierende Features findet. Andererseits soll untersucht werden, wie die Verteilung und die Anzahl der Features in dem Bildausschnitt der Kameras die Genauigkeit beeinflusst.

In Abschnitt 4.1 werden daher die Art und Geometrie der Texturen für die Experimente erläutert. Abschnitt 4.2 beschreibt die Konzeption der Bewegungen. Der Roboterarm führt mit den montierten Stereokameras exakt definierte Translationen und Rotationen durch. Der Algorithmus zur Berechnung der visuellen Odometrie soll bei dieser Arbeit für den Einsatz auf dem Krabblar untersucht werden. Daher werden Art und Umfang der ausgeführten Bewegungen so gewählt, dass sie in etwa denen des Krabblers entsprechen. In Abschnitt 4.3 wird der Ablauf der Experimente erläutert. Für eine exakte Auswertung der Daten werden für jede Bewegung der Stereokameras ein Video auf der Startposition und eines auf der Endposition aufgenommen. Die Kameras sind dabei pro Experiment auf unterschiedliche Variationen der Textur gerichtet. Durch die Konzeption der Experimente werden bestimmte Ergebnisse der Genauigkeitsschätzungen erwartet. Diese werden in Abschnitt 4.4 behandelt.

4.1 Verwendete Texturen

Damit der Egomotionalgorithmus eine Bewegung zwischen zwei Bildern schätzen kann, müssen korrespondierende Features in den jeweils aufeinanderfolgenden linken Kamerabildern gefunden werden. Für die Genauigkeit der Bewegungsschätzung ist die Anzahl, die Art und die Verteilung der gefundenen Features von Bedeutung. Die Features erkennt der Algorithmus in der Textur der Objekte des aufgezeichneten

Bildausschnitts. Daher werden bei den Experimenten mit Hilfe der Texturen die detektierten Features beeinflusst. Dabei sollen einerseits maximale Genauigkeiten der Bewegungsschätzung erzielt werden, andererseits können Fehler anhand der Größe der texturierten Fläche oder der Anzahl der gefundenen Features provoziert werden. Je mehr korrespondierende Features detektiert werden, desto zuverlässiger wird die Bewegungsschätzung. Dies setzt allerdings voraus, dass die verschiedenen Features voneinander unterschieden werden können. Weist die Textur viele sich wiederholende Strukturen auf, werden leichter falsche Korrespondenzen detektiert und die Schätzung wird trotz vieler Features ungenau. Diese sind bei künstlich geschaffenen Texturen, wie z.B. einem Gitter, häufig zu finden. Im Gegensatz dazu haben natürliche Texturen kaum identische Strukturen.

Um die richtige Textur für die Experimente zu finden, wurden eine künstliche Gittertextur, die Textur eines Plastikbodens und eine natürliche Korktextur getestet. Tabelle 4.1 zeigt die von dem Egomotionalgorithmus gefundene Anzahl von korrespondierenden Features.

	Abstand 60cm			Abstand 30cm		
	corners			corners		
	det.	corr.	corr./det.	det.	corr.	corr./det.
Gitter	700	500	0,714	600	500	0,833
Kork	700	600	0,857	600	500	0,833
Plastik	2000	1000	0,500	250	100	0,400

Tabelle 4.1 Vergleich von verschiedenen Texturen

Dabei wurden die Stereokameras in 60cm und 30cm Höhe über die Texturen gehalten und die ungefähre Anzahl der gefundenen Corners „corners det.“ und der korrespondierenden Corners „corners corr.“ notiert. Da die Korktextur am meisten korrespondierende Corners aufweist und im Gegensatz zu den anderen beiden aus natürlichen Strukturen besteht, eignet sich diese für die Experimente am besten. Eine Textur, bei der nur wenige Corners gefunden werden, kann später in Abschnitt 5.2 bei der Auswertung der Daten durch Anpassung der Parameter zur Featuredetektion simuliert werden. Somit sind weniger Experimente notwendig.

Um die Verteilung der Features in dem Bildausschnitt zu steuern, werden von der Korktextur unterschiedlich große Quadrate angefertigt. Die detektierten Features sollen nur innerhalb dieser Texturen gefunden werden. Für die ersten Versuche sollen sich die Features auf eine rechteckige Fläche von 8x8cm, 20x20cm und 30x30cm

konzentrieren. In dem letzten Experiment soll die Korktextur den gesamten Bildausschnitt der Stereokameras füllen. Die korrespondierenden Features werden so

(a) 8x8cm



(b) 20x20cm



(c) 30x30cm



(d) komplette Fläche



Abbildung 4.1 unterschiedliche Größen der texturierten Korkflächen

lückenlos über den gesamten Bildausschnitt verteilt gefunden. Abbildung 4.1 zeigt das linke Stereokamerabild der jeweiligen Texturen. Durch die intrinsische Kalibrierung der Kameras kann angenommen werden, dass das rektifizierte Bild der Kameras unvermerkt vorliegt und es keinen Unterschied für die Genauigkeit macht, ob sich ein Feature am Rand oder in der Mitte des Bildes befindet. Anhand der quadratischen Texturen und der rechteckigen Form des Bildausschnitts kann problemlos eine Abdeckung der Bildfläche durch Korktexturen errechnet werden. Die Pixelwerte der Texturen in Tabelle 4.2 wurden aus den Stereokamerabildern entnommen.

	Abstand 65cm		Abstand 105cm	
Korktextur	Größe Textur	Fläche	Größe Textur	Fläche
8cm	88px*88px	~1%	-	-
20cm	220px*220px	~10%	-	-
30cm	326px*326px	~20%	207px*207px	~10%
komplett	814px*316px	100%	814px*316px	100%

Tabelle 4.2 Fläche der unterschiedlichen Korktexturen im digitalen Bild

4.2 Art der Bewegungen

Neben den Texturen spielt auch die Art und der Umfang der Bewegung bei den Experimenten eine Rolle. Da die Aufnahme der Tests für alle möglichen Translationen und Rotationen in x-, y- und z-Richtung zeitlich zu aufwändig ist, werden in dieser Arbeit nur die Bewegungen entlang und senkrecht zur optischen Kameraachse getestet. Vom Kamerakoordinatensystem ausgehend entspricht das Translationen in x- und z-Richtung, sowie Rotationen um die y- und z-Achse. Eine Translation in y-Richtung ist vom Bewegungsverlauf analog zu der x-Richtung, eine Rotation um die x-Achse analog zu der um die y-Achse.

Der Egomotionalgorithmus wird auf dem Krabbler in der Regel mit einer Frequenz von 5 Hz genutzt. Es werden alle 0,2 Sekunden Bilder für den Algorithmus aufgenommen und relative Translation und Rotation zwischen den aufeinanderfolgenden Bildpaaren gerechnet. Für diese Arbeit ist nur die relative Bewegung von einem Zeitintervall zum darauf folgenden von Bedeutung. Da die Bewegungen des Krabblers zwischen den Bildpaaren sehr klein sind, werden bei den Testreihen Translationen von 10mm, 20mm und 50mm und Rotationen von 2° und 10° durchgeführt.

Die Tests werden mit den Abständen der Stereokameras zu den Texturen von 650mm und 1050mm durchgeführt. Somit kann in der anschließenden Auswertung auch ein Zusammenhang zwischen der Genauigkeit der Bewegungsschätzung und dem Abstand zu den Texturen hergestellt werden.

4.3 Datenaufnahme

Für die Datenaufzeichnung werden die in Kapitel 4.2 beschriebenen Bewegungen in die Steuerungssoftware des KR-16 einprogrammiert. Für Translationen wird eine Trajektorie des TCP entlang der jeweiligen Kameraachse bestimmt. Die Rotationen werden über eine Drehung des TCP um die Kameraachsen entsprechend den verschiedenen Gradzahlen festgelegt. Die Bewegungen werden statisch aufgezeichnet, pro Bewegung ein Video mit jeweils 5 Sekunden und 5 fps auf der Startposition und eines auf der Endposition. Somit kann die relative Bewegung für ein Zeitintervall in der Auswertung redundant berechnet werden, ohne dass Ungenauigkeiten des Roboterarms durch wiederholtes Anfahren einer Position das Ergebnis beeinflussen.

Für alle Translationen mit einem Abstand zwischen Kameras und texturierter Fläche von 650mm werden die Korktexturen mit der vollen Bildabdeckung und mit den Größen von 8x8cm, 20x20cm und 30x30cm durch gewechselt. Die Rotationen werden nur mit den Texturen 30x30cm und der vollen Bildabdeckung durchgeführt, da mit den kleineren texturierten Flächen der Algorithmus zu schlechte Ergebnisse schätzen würde. Bei dem Abstand von 1050mm wird für die Translationen und Rotationen die volle Bildabdeckung verwendet. Daraus resultieren 42 Experimente, die sich aus drei Translationen entlang jeweils zweier Achsen, zwei Rotationen um jeweils zwei Achsen, vier verschiedenen Texturformen und zwei unterschiedlichen Abständen der Kameras zu den Texturen zusammensetzen.

Bei den Versuchen wird darauf geachtet, dass jeweils die gleichen Lichtverhältnisse vorherrschen, damit die Belichtung und Verschlusszeit der Kameras immer gleich ist. Um ein möglichst genaues Ergebnis bei der Auswertung erzielen zu können, werden diese so gewählt, dass die visuelle Odometrie zwischen den Bildern ein Maximum von korrespondierenden Features erkennt. Dabei wird auch darauf geachtet, dass nur Features auf den vorgesehenen texturierten Flächen gefunden werden, und nicht im restlichen Bildausschnitt.

4.4 Erwartete Ergebnisse

Die Experimente wurden entsprechend konzipiert, dass die Bewegungsschätzung so beeinflusst wird, damit die verschiedenen Experimente bei der Auswertung unterschiedliche Genauigkeiten vorzeigen.

Bei den translatorischen sowie rotatorischen Bewegungen werden Abweichungen in der Genauigkeit abhängig von der Bewegungsrichtung erwartet. Durch die Berechnungsart der Bewegungsschätzung macht es einen Unterschied, ob sich die Kameras parallel zur Bildebene bewegen, oder auf sie zu. Das Gleiche gilt für Rotationen. Drehungen um die optische Achse haben andere geometrische Effekte, als Rotationen um die x- oder y-Achse.

Die Größe der durchgeführten Translationen und Rotationen sollte keinen Einfluss auf die Genauigkeit der Bewegungsschätzung haben. Bei der Maximaltranslation von 50mm und Maximalrotation von 10° überlappen sich beide Bilder von Start- und Endposition immer noch sehr stark. Des Weiteren ändert sich der Blickwinkel zu den Texturen nur unwesentlich. Daher können die korrespondierenden Features immer noch problemlos detektiert werden.

Es wird davon ausgegangen, dass die Anzahl der gefundenen Features einen großen Einfluss auf die Genauigkeit der Schätzung hat. Aus den bisherigen Erfahrungswerten sind ca. 100 korrespondierende Features über das ganze Bild verteilt optimal für eine gute Bewegungsschätzung. Bei 10 bis 20 Features muss die Genauigkeit der Schätzung deutlich schlechter werden. Die Corners reichen dann für eine zuverlässige Bewegungsschätzung nicht mehr aus.

Ein weiterer großer Einfluss auf die Genauigkeit wird von der Größe der texturierten Fläche erwartet. Je kleiner die texturierte Fläche im aufgenommenen Bild ist, desto mehr werden die Features auf einen kleinen Raum des Bildes konzentriert. Es wird vermutet, dass die Genauigkeit der Bewegungsschätzung mit kleineren texturierten Flächen abnimmt. Vor allem bei der Texturflächengröße 8x8cm müssten die Ergebnisse sehr schlecht werden.

Bei einer Vergrößerung des Abstands der Stereokameras zu den aufgenommenen Texturen muss die Genauigkeit der Schätzung abnehmen. Dieser Genauigkeitsunterschied sollte vor allem bei den texturierten Flächen, die das ganze Bild abdecken, bemerkbar sein. Trotz gleicher Flächenabdeckung im Bild muss die Genauigkeit der translatorischen Schätzungen abnehmen, da die Koordinaten der markanten Punkte in der Textur ungenauer rekonstruiert werden [2, S. 55-56]. Auf den rotatorischen Anteil der Schätzung sollte eine Abstandsvergrößerung keinen Einfluss haben.

Der Vergleich der Fehlerschätzung mit den tatsächlichen Fehlern müsste zeigen, dass der reale Fehler nach der KR-16 Messung immer kleiner ist, da ein maximaler Fehler geschätzt wird.

Kapitel 5

Rechnerische Auswertung der Ergebnisse

Dieses Kapitel beschreibt die rechnerische Auswertung der aufgenommenen Daten der Experimente aus Kapitel 4.

Abschnitt 5.1 erklärt die Formelzeichen, die in den Abschnitten des 5. Kapitels verwendet werden. Abschnitt 5.2 beschreibt, wie aus den aufgenommenen Videos mit Hilfe der Stereolibrary die Bewegungsschätzung der Stereokameras berechnet wird. Diese werden anschließend mit den tatsächlichen Bewegungen des Roboterarms mit Hilfe von Matlab in Abschnitt 5.3 gegenübergestellt. Abschnitt 5.4 geht auf die Aufbereitung der Daten ein, damit diese in Kapitel 6 interpretiert werden können.

5.1 Erklärung der Formelzeichen

Dieser Abschnitt erklärt die verwendeten Notationen für die folgenden Abschnitte des 5. Kapitels.

Allgemeine Formelzeichen:

T_{XY}^Z	Translationsvektor von Koordinatensystem X in Koordinatensystem Y im Referenzsystem Z
R_{XY}	Rotationsmatrix von Koordinatensystem X in Koordinatensystem Y
$\langle axis, angle \rangle_{XY}$	axis-angle-Darstellung der Rotation von Koordinatensystem X in Koordinatensystem Y
$\langle euler \rangle_{XY}$	Euler-Darstellung der Rotation von Koordinatensystem X in Koordinatensystem Y
σ	Standardabweichung

Indizes:

W	Weltkoordinatensystem
tcp	TCP-Koordinatensystem
$tcp1$	TCP-Koordinatensystem im Startpunkt

<i>tcp2</i>	TCP-Koordinatensystem im Endpunkt
<i>K</i>	Kamerakoordinatensystem
<i>K1</i>	Kamerakoordinatensystem im Startpunkt
<i>K2</i>	Kamerakoordinatensystem im Endpunkt
<i>err</i>	geschätzte Fehler der visuellen Odometrie
<i>res</i>	Residuen
<i>n</i>	Index der geschätzten Egomotion pro Bewegung

5.2 Berechnung der Bewegungsschätzung

Um eine Bewegungsschätzung der Kameras für die Experimente zu bekommen, stehen vier verschiedene Kommandozeilenprogramme von Dr. Heiko Hirschmüller zur Verfügung. Diese basieren auf dem Algorithmus der visuellen Odometrie. Das Programm „avitool“ extrahiert aus den Videos aus Abschnitt 4.3 ein PGM-Stereobildpaar pro Sekunde. Danach werden die Bilder mit dem Programm „rectify“ rektifiziert. Über „stereo“ wird aus den aufeinanderfolgenden Bildpaaren ein Disparitätsbild erstellt. „motion“ berechnet über die rektifizierten Bilder und dem Disparitätsbild die Bewegungsschätzung über den Egomotionalgorithmus. Da am DLR als Betriebssystem Linux eingesetzt wird, wird über eine in der Konsole ausführbare Bashdatei eine Routine geschrieben, die zuerst die Bilder aus den Videos extrahiert. Aufgrund der Rechenzeit wird pro Sekunde Videolaufzeit nur ein Bild extrahiert. Danach liegen pro Bewegung fünf Stereobildpaare von der Startposition und fünf Paare von der Endposition vor. Die Routine berechnet anschließend pro Bild von Start- und Endposition eine Bewegungsschätzung.

Daraus resultieren 25 geschätzte Bewegungen für eine einzige Bewegung aus dem jeweiligen Experiment. Diese Redundanz erhöht die Genauigkeit bei der Auswertung. Für die Programme „rectify“, „stereo“ und „motion“ wird die Parameterdatei (siehe Abschnitt 3.4) benötigt. Neben den intrinsischen und extrinsischen Kalibrierdaten der Stereokameras sind dort auch die notwendigen Einstellungen für den Egomotionalgorithmus festgelegt. Dazu gehört unter anderem der Parameter *t* der Corner Detection. Er gibt an, wie markant ein Corner sein muss, um detektiert und zur Motionberechnung verwendet zu werden. Die Werte werden für jede Bewegung so gesetzt, dass für die Berechnung circa 100 korrespondierende Corners verwendet werden. Bei Bewegungen, wo das Bild komplett mit Textur ausgefüllt ist, müssen die Schwellenwerte höher gesetzt werden, als bei Bewegungen mit weniger

Texturabdeckung. Der Schwellenwert wird auch dazu benutzt, um Datensätze zu erzeugen, die schlechte Texturen simulieren sollen. Dafür werden neue Bewegungsschätzungen durchgeführt, in denen nur 50, 30 und 10 korrespondierende Corners für die Berechnung verwendet werden.

Die Schätzung wird als relative Translation in x-, y- und z-Richtung $T_{K1 K2}^{K1}$ und einer relativen Rotationsmatrix $R_{K1 K2}$ in eine Textdatei ausgegeben. Zusätzlich wird pro Schätzung ein von dem Algorithmus geschätzter Fehler für Translation $T_{err K1 K2}^{K1}$ und Rotation $R_{err K1 K2}$ mit berechnet. In einer zweiten Textdatei werden die Koordinaten der verwendeten Corners als x- und y- Wert, bezogen auf das Bildkoordinatensystem der aufgenommenen Stereobilder, gespeichert. Anhand dieser Information kann später die Verteilung der Corners im Bildausschnitt berechnet werden.

5.3 Genauigkeitsvergleich mit Referenzdaten

Damit die geschätzten Bewegungen mit den tatsächlichen Bewegungen verglichen werden können, werden die Daten des KR-16 und die erstellten Textdateien mit der berechneten Motion in Matlab eingelesen.

5.3.1 Transformation der KR-16 Daten

Die Translationen und Rotationen des TCP von Anfangspunkt $T_{W tcp1}^W$ und $R_{tcp1 W}$, sowie von Endpunkt $T_{W tcp2}^W$ und $R_{tcp2 W}$ jeder Bewegung werden aus dem Steuerungsprogramm ausgelesen und in Matlab gespeichert. Dabei ist im Gegensatz zu der Egomotion angegeben, wie die Translation und Rotation des TCP- zu dem Welt-Frame auf dem jeweiligen Punkt ist. Damit die Bewegung des Arms mit der der Egomotion verglichen werden kann, muss zuerst die relative Bewegung $T_{tcp1 tcp2}^{tcp1}$ und $R_{tcp1 tcp2}$ des TCP von Start- zu Endpunkt berechnet werden (Formel (5.1) und (5.2)).

$$R_{tcp1 tcp2} = R_{tcp2 W}^T \cdot R_{tcp1 W} \quad (5.1)$$

$$T_{tcp1 tcp2}^{tcp1} = R_{tcp1 W}^T \cdot (T_{W tcp2}^W - T_{W tcp1}^W) \quad (5.2)$$

Danach wird mit Hilfe der Parameter $T_{K tcp}^K$ und $R_{tcp K}$ aus der Hand – Auge Kalibrierung (siehe Abschnitt 3.4) mit den Formeln (5.3) und (5.4) die Bewegung $T_{K1 K2}^{K1}$ und $R_{K1 K2}$ der Stereokameras von Start- zu Endpunkt berechnet.

$$R_{K1 K2} = R_{tcp K} \cdot R_{tcp1 tcp2} \cdot R_{tcp K}^T \quad (5.3)$$

$$T_{K1 K2}^{K1} = T_{K tcp}^K + R_{tcp K} \cdot T_{tcp1 tcp2}^{tcp1} - R_{K1 K2}^T \cdot T_{K tcp}^K \quad (5.4)$$

Jetzt liegt die Referenzbewegung der Stereokameras im gleichen Bezugssystem vor, wie die zu überprüfenden Bewegungen der Egomotion.

5.3.2 Auswertung der motion.txt Dateien

Als zweiter Schritt werden die in Abschnitt 5.3.1 erstellten Textdateien für jede Bewegungsschätzung der Egomotion in Matlab eingelesen. Die Bewegung der Egomotion liegt als Verschiebung $T_{n K1 K2}^{K1}$ und $R_{n K1 K2}$ des ersten Kameraframes zu dem zweiten Kameraframe vor. n steht für den Index der geschätzten Egomotion pro Bewegung. Da pro Bewegung 25 Bilder gerechnet werden gilt $n = 1 \dots 25$. Diese 25 Schätzungen werden gemittelt. Zuerst wird mit Formel (5.5) aus den Translationen der arithmetische Mittelwert gebildet. Um die Rotationen zu mitteln, werden die Rotationsmatrizen $R_{n K1 K2}$ in axis-angle-Darstellungen $\langle axis, angle \rangle_{n K1 K2}$ umgewandelt und daraus mit Formel (5.6) der arithmetische Mittelwert $\overline{\langle axis, angle \rangle}_{K1 K2}$ erstellt. Durch die spezielle Darstellungsform der axis-angle Darstellung ist es möglich, die Winkel pro Achse arithmetisch zu mitteln, ohne dabei den Bezug zu den anderen zwei Achsen zu verfälschen.

$$\bar{T}_{K1 K2}^{K1} = \frac{\sum_1^n T_{n K1 K2}^{K1}}{n} \quad (5.5)$$

$$\overline{\langle axis, angle \rangle}_{K1 K2} = \frac{\sum_1^n \langle axis, angle \rangle_{n K1 K2}}{n} \quad (5.6)$$

Nach Umwandlung der axis-angle-Darstellung in eine Rotationsmatrix liegt die Bewegungsschätzung des Egomotionalgorithmus jetzt als gemittelte Translation $\bar{T}_{K1 K2}^{K1}$ und Rotation $\bar{R}_{K1 K2}$ vor.

Analog zu der berechneten Bewegung der Stereokameras werden aus den Textdateien die von der visuellen Odometrie geschätzten Fehler der Bewegung $T_{n_{err_{K1K2}}}^{K1}$ und $R_{n_{err_{K1K2}}}$ ausgelesen und gemittelt. Aus den Formeln (5.7) und (5.8) ergeben sich $\bar{T}_{err_{K1K2}}^{K1}$ und $\overline{\langle axis, angle \rangle}_{err_{K1K2}}$.

$$\bar{T}_{err_{K1K2}}^{K1} = \frac{\sum_1^n T_{n_{err_{K1K2}}}^{K1}}{n} \quad (5.7)$$

$$\overline{\langle axis, angle \rangle}_{err_{K1K2}} = \frac{\sum_1^n \langle axis, angle \rangle_{err_{n_{K1K2}}}}{n} \quad (5.8)$$

5.3.3 Auswertung der corner.txt Dateien

Zu jeder geschätzten Bewegung werden von dem Egomotionalgorithmus die Bildkoordinaten der dafür verwendeten Corners in einer Textdatei gespeichert. Diese Dateien werden auch in Matlab eingelesen, um den Bezug der Genauigkeit der Bewegungsschätzung zu der Verteilung der verwendeten Corners herzustellen. Dazu wird von jedem Cornerpunkt die Strecke zu dem Schwerpunkt aller Cornerpunkte berechnet. Die Strecken werden arithmetisch gemittelt, damit pro Bewegungsschätzung ein Abstandswert \bar{d} der Corners ausgegeben wird.

5.3.4 Berechnung der Genauigkeiten

Um eine Aussage über die Genauigkeit der berechneten Egomotion zu treffen, wird der mittlere Fehler zu den in den $K1$ -frame transformierten KR-16- Werten aus Formel (5.3) und (5.4) berechnet. Im weiteren Verlauf der Arbeit werden diese Fehler als Residuen der Translation T_{res} und der Rotation $\langle axis, angle \rangle_{res}$ bzw. R_{res} bezeichnet.

$$T_{res} = \frac{\sum_1^n \sqrt{(T_{K1K2}^{K1} - T_{n_{K1K2}}^{K1})^2}}{n} \quad (5.9)$$

$$\langle axis, angle \rangle_{res} = \frac{\sum_1^n \sqrt{(\langle axis, angle \rangle_{K1K2} - \langle axis, angle \rangle_{n_{K1K2}})^2}}{n} \quad (5.10)$$

Außerdem werden über die Formel (5.11) bis (5.14) noch die Standardabweichung von Translation und Rotation der visuellen Odometrie zu dem jeweiligen Mittelwert $\sigma_{\bar{T}}$ bzw. $\sigma_{\langle axis, angle \rangle}$ und zu dem KR-16 Sollwert σ_T bzw. $\sigma_{\langle axis, angle \rangle}$ berechnet.

Standardabweichung der visuellen Odometrie zu dem jeweiligen Mittelwert

$$\sigma_{\bar{T}} = \sqrt{\frac{\sum_1^n \sqrt{(\bar{T}_{K1 K2}^{K1} - T_{n K1 K2}^{K1})^2}}{n-1}} \quad (5.11)$$

$$\sigma_{\langle axis, angle \rangle} = \sqrt{\frac{\sum_1^n \sqrt{(\langle axis, angle \rangle_{K1 K2} - \langle axis, angle \rangle_{n K1 K2})^2}}{n-1}} \quad (5.12)$$

Standardabweichung der visuellen Odometrie zu dem jeweiligen KR-16 Sollwert

$$\sigma_T = \sqrt{\frac{\sum_1^n \sqrt{(T_{K1 K2}^{K1} - T_{n K1 K2}^{K1})^2}}{n-1}} \quad (5.13)$$

$$\sigma_{\langle axis, angle \rangle} = \sqrt{\frac{\sum_1^n \sqrt{(\langle axis, angle \rangle_{K1 K2} - \langle axis, angle \rangle_{n K1 K2})^2}}{n-1}} \quad (5.14)$$

Für die weitere Auswertung in den folgenden Abschnitten wird die axis-angle-Darstellung ($\langle axis, angle \rangle$) der Winkel in Eulerwinkel ($\langle euler \rangle$) umgerechnet.

5.4 Grafische Aufbereitung in Excel

Die Ergebnisse aus Matlab werden in eine Excel-Liste (siehe Anhang) übertragen. Für eine getestete Bewegung werden jeweils die Residuen T_{res} bzw. $\langle euler \rangle_{res}$, der geschätzte translatorische Fehler der visuellen Odometrie $\bar{T}_{err_{K1 K2}}^{K1}$, die Standardabweichung zum Mittelwert $\sigma_{\bar{T}}$ bzw. $\sigma_{\langle euler \rangle}$, die Standardabweichung zum Soll-Wert σ_T bzw. $\sigma_{\langle euler \rangle}$, die Anzahl der gefundenen Corners und der mittlere Abstand der Corners \bar{d} zum Schwerpunkt gegenübergestellt.

Die Residuen dienen als Maß für die Genauigkeit der Schätzung. Die Standardabweichung der visuellen Odometrie zum Mittelwert zeigt die Eindeutigkeit der Schätzung. Die Standardabweichung zum Soll-Wert kann in Verbindung mit den Residuen Aufschluss darüber geben, wie eindeutig die Residuen berechnet wurden. Die Anzahl und der mittlere Abstand der gefundenen Corners dienen als Kontrolle in der Datentabelle. Somit können Anzahl der für die Berechnung verwendeten Corners und Änderungen der Größe der texturierten Fläche in der Tabelle numerisch sichtbar gemacht werden. Mit Hilfe von Charts werden ausgewählte Daten in Kapitel 6 grafisch miteinander verglichen.

Kapitel 6

Visualisierung und Interpretation der Genauigkeiten

In diesem Kapitel werden die durchgeführten Experimente auf verschiedene Aspekte ausgewertet. Grundlage dafür sind die aufbereiteten Daten aus Kapitel 5. Diese werden mit verschiedenen Grafiken visualisiert, um die Ergebnisse besser interpretieren zu können. Ziel ist es dabei einerseits, einen Zusammenhang zwischen den korrespondierenden Features im Stereokamerabild und der Genauigkeit der geschätzten Bewegung des Egomotionalgorithmus herzustellen. Andererseits soll die Fehlerschätzung des Algorithmus überprüft werden.

Als Maß für den Fehler in den folgenden Abschnitten werden die Residuen aus Abschnitt 5.3.4 verwendet. Sie beschreiben den mittleren quadratischen Fehler der visuellen Odometrie bezogen auf die Messungen des KR-16. Die beschriebenen Koordinatenachsen beziehen sich auf das Stereokamerasystem, welches in Abschnitt 3.3 erläutert wurde. Demnach geht die x-Achse nach rechts, die y-Achse zeigt nach unten und die z-Achse ist die optische Achse der linken Kamera bezogen auf die Blickrichtung der Kamera.

Der erste Aspekt der Genauigkeitsuntersuchung ist der Vergleich der tatsächlich durchgeführten Bewegung des KR-16 mit der berechneten Bewegung des Egomotionalgorithmus in Abschnitt 6.1. Dabei interessiert die Genauigkeit des Algorithmus in Bezug zur Bewegungsgröße. Danach wird die Abhängigkeit der Genauigkeit von den Eigenschaften der korrespondierenden Corners untersucht. Abschnitt 6.2 geht auf die Anzahl der verwendeten Corners für eine Bewegungsschätzung ein. In Abschnitt 6.3 wird die Abhängigkeit von der Verteilung der korrespondierenden Corners im aufgenommenen Bild dargestellt. Abschnitt 6.4 zeigt, wie die Residuen von dem Abstand der Stereokameras zu der aufgenommenen Szene abhängen. Abschnitt 6.5 vergleicht die tatsächlichen Fehler der Bewegungsschätzung mit der Fehlerschätzung des Egomotionalgorithmus mit dem Ziel, die Annahmen der Fehlerschätzung zu verifizieren. Im letzten Abschnitt 6.6 wird das Fehlerverhalten des Egomotionalgorithmus aus den einzelnen Untersuchungen zusammengefasst.

6.1 Erreichte Genauigkeit in Abhängigkeit der Größe der Bewegung

Dieser Abschnitt soll die maximal erreichbare Genauigkeit des Egomotionalgorithmus aufzeigen. Dabei wird auch der Einfluss der Bewegungsgröße auf die Genauigkeit untersucht. Die Abbildungen 6.1 bis 6.4 vergleichen dafür die Residuen zwischen der Bewegung des KR-16 und der berechneten visuellen Odometrie. Es sollen nur die Daten mit höchster Genauigkeit untersucht werden. Daher werden die Berechnungen der Tests verwendet, bei denen der gesamte Bildausschnitt der Stereokameras mit einer Korktextur ausgefüllt ist. Für die Schätzung der Bewegungen der Egomotion werden jeweils ca. 100 Features verwendet und der Abstand der Kameras zu den texturierten Flächen liegt bei 650mm.

6.1.1 Translatorische Bewegungen

Abbildung 6.1 zeigt die Residuen der Translationen der Stereokameras in z-Richtung, auf die ausgenommene Szene zu. Der maximale Fehler beträgt 0,53mm bei 50mm Translation. Die restlichen Fehler schwanken zwischen 0,05mm und 0,3mm. Der maximale Fehler in der Rotation liegt bei 0,035°. Generell zeigt die Abbildung, je größer

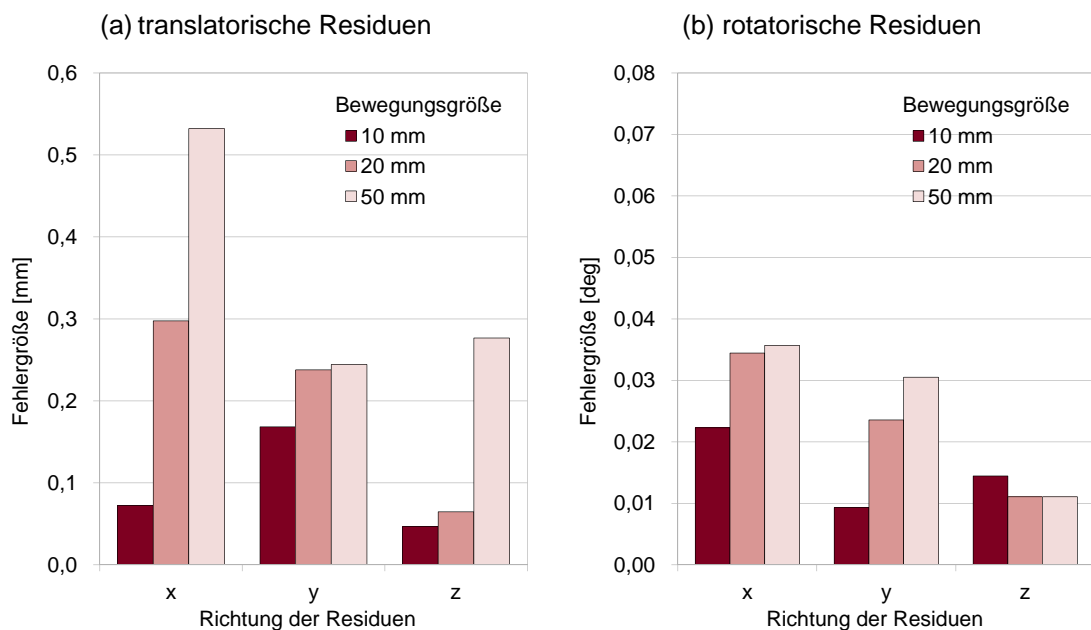


Abbildung 6.1 Translation in z-Richtung, Abhängigkeit der Residuen von Bewegungsgröße

die Bewegung ist, desto ungenauer wird die Schätzung in x-, y- und z-Richtung. Dies gilt für die translatorischen und rotatorischen Fehler. Ausnahme sind die rotatorischen Residuen in z-Richtung bei den Bewegungen 20mm und 50mm. Hier zeigt sich, dass die Schätzung für die Anteile der Rotationen um die z-Achse bei einer Bewegung generell sehr genau zu bestimmen ist. Die Features legen in den Bildausschnitten den größten Weg zurück, daher lässt sich eine Bewegung über diese Features sehr genau bestimmen. Bei rotatorischen Anteilen um die z-Achse wird die gesamte Strecke der Rotation in der Bildebene zurück gelegt. Im Gegensatz dazu legen die Features bei rotatorischen Anteilen um die x- oder y-Achse den größten Teil der von der Rotation verursachten Strecke senkrecht zur Bildebene zurück.

Abbildung 6.2 zeigt die Residuen einer Translation der Kameras in x-Richtung, parallel zu der texturierten Fläche. Der maximale Fehler beläuft sich auf 0,26mm bei 50mm Translation. Die restlichen Fehler schwanken hauptsächlich zwischen 0,1mm und 0,2mm. Der maximale Fehler in der Rotation liegt bei 0,07°. Die Größe der Bewegung

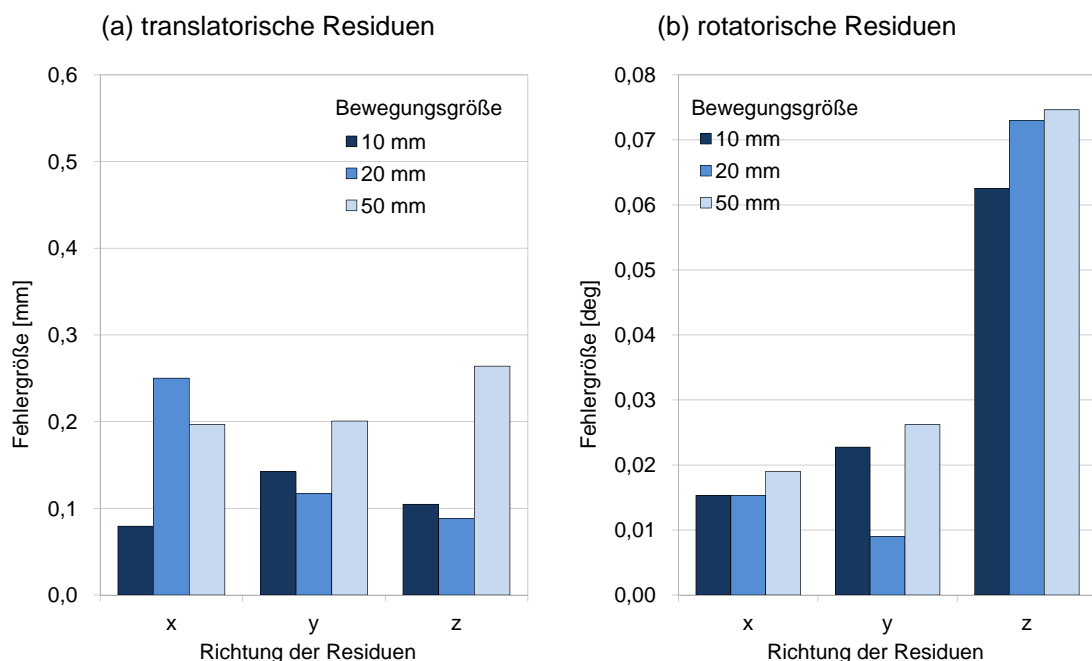


Abbildung 6.2 Translation in x-Richtung, Abhängigkeit der Residuen von Bewegungsgröße

zeigt bei einer Translation in x-Richtung keinen direkten Einfluss auf die Genauigkeit der Schätzung, wie in Abschnitt 4.4 erwartet. Die rotatorischen Fehler sind um die z-Achse am größten, was für alle drei Bewegungsgrößen gilt. Es wurde erwartet, dass die Residuen auch hier sehr klein sind. Jede Bewegungsgröße entspricht einer von den anderen Datensätzen unabhängigen Aufnahme der Daten. Daher kann ein Fehler

in der Aufnahme ausgeschlossen werden. Anhand der aufgeführten Standardabweichungen der visuellen Odometrie zu dem Mittelwert der 25 geschätzten Bewegungen in der Datentabelle kann man dieses Verhalten für die drei Bewegungsgrößen genauer untersuchen (vgl. Anhang, A.1). Die Standardabweichung ist für den rotatorischen Fehler in z-Richtung bei einer Translation in x-Richtung für alle Bewegungsgrößen sehr klein. Der Algorithmus schätzt also die Bewegung sehr eindeutig, da die Werte der Schätzung nur gering schwanken. Die Standardabweichungen für den Fehler in x- und y-Richtung sind im Vergleich dazu deutlich größer. In Abschnitt 6.2.1 wird auf diese Problematik noch einmal eingegangen und weitere Datensätze zeigen das gleiche Verhalten. Daraus lässt sich schließen, dass ein systematischer Fehler vorliegt. Ob sich dieser aus der Schätzung oder einem Fehler des Roboterarms ergibt, müsste mit weiteren Experimenten untersucht werden.

Im direkten Vergleich der Residuen der beiden Bewegungsrichtungen sind die translatorischen Fehler bei einer Translation in x-Richtung durchschnittlich etwas kleiner, als bei einer Bewegung in z-Richtung. Das hat zwei Gründe. Einerseits ist eine Bewegung parallel zur Bildebene generell besser zu detektieren, als eine entlang der optischen Achse. Hier findet die ganze Strecke der Bewegung in der Bildebene statt. Bei einer Translation entlang der optischen Achse bewegen sich die Features senkrecht zur Bildebene und ändern nur ihre Entfernung zu den Kameras. Die beschriebene Strecke der korrespondierenden Features ist also bei einer Translation in z-Richtung bei gleicher Bewegungsgröße kleiner. Ein weiterer Grund für die unterschiedlich großen Residuen der beiden Translationsrichtungen ist, dass bei einer translatorischen Bewegung in z-Richtung, also auf die Szene zu, eine Skalierung der Bildinhalte auftritt. Diese Skalierung wird allerdings nicht von dem Featurematching-Verfahren, welches in Abschnitt 2.3.1 beschrieben wird, berücksichtigt. Der Suchbereich, der um den markanten Corner-Punkt festgelegt wird, bleibt bei beiden Bildern gleich groß. Somit ändern sich die Inhalte der Suchbereiche, die miteinander verglichen werden. Je größer die Bewegung in Richtung der Szene ist, desto unterschiedlicher sind die Inhalte. Korrespondenzen können dann nicht mehr exakt gefunden werden und der Egomotionalgorithmus liefert ungenauere Positionsschätzungen. Aus diesem Grund steigt der Fehler der Egomotionschätzung mit zunehmender Bewegungsgröße in z-Richtung an. Bei einer Bewegung in x-Richtung findet keine Skalierung der Features statt, darum ist keine Abhängigkeit von der Bewegungsgröße erkennbar.

6.1.2 Rotatorische Bewegungen

Die Residuen der Rotationen um die z-Achse werden in Abbildung 6.3 visualisiert. Dabei tritt ein maximaler rotatorischer Fehler von $0,10^\circ$ auf. Der maximale translatorische Fehler liegt bei $0,87\text{mm}$. Auch hier zeigt sich wieder, dass die translatorischen und rotatorischen Fehler sehr stark ansteigen, wenn die durchgeführte Rotation vergrößert wird. Das ist mit einer nicht vorhandenen Rotationsinvarianz des Featurematching-Verfahrens zu erklären. Drehen sich die Kameras um die z-Achse, drehen sich auch die Inhalte der Bilder relativ zueinander gesehen. Bei dem Featurematching-Verfahren bleibt der Suchbereich um die signifikanten Corners aber wieder gleich und rotiert nicht mit. Somit ändern sich wieder die Inhalte der Suchbereiche und die Schätzungen werden fehlerhaft. Der rotatorische und translatorische Fehler in z-Richtung ist am kleinsten, vor allem bei einer Rotation um 10° . Hier zeigt sich wieder, dass die z-Richtung der rotatorischen Komponente einer Bewegung sehr genau bestimmt werden kann.

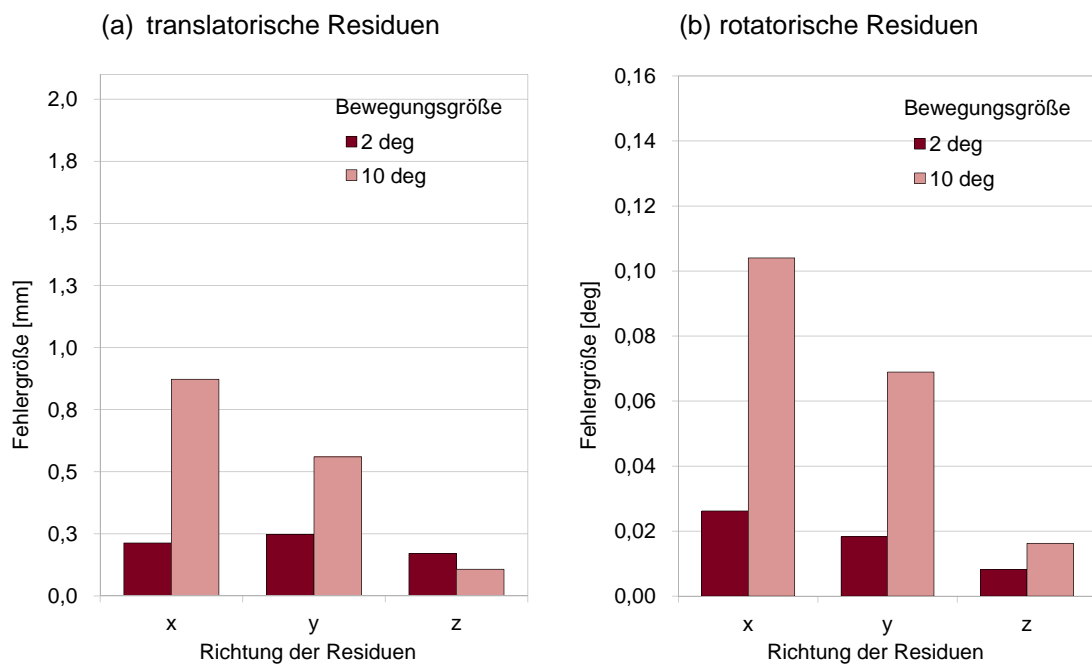


Abbildung 6.3 Rotation um z-Achse, Abhängigkeit Residuen von Bewegungsgröße

Abbildung 6.4 visualisiert eine Rotation um die y-Achse. Der maximale rotatorische Fehler liegt bei $0,16^\circ$. Der maximale translatorische Fehler bei $2,0\text{mm}$. Wie in Abbildung 6.3 ist auch hier ein Anstieg der Residuen mit steigender Bewegungsgröße zu erkennen. Allerdings sind die Differenzen der Residuen deutlich kleiner. Bei

Rotationen um die y-Achse rotieren die Features im Bild jedoch nicht mit, deswegen kann die fehlende Rotationsinvarianz des Featurematching-Verfahrens nicht die Ursache für den Anstieg des Fehlers sein.

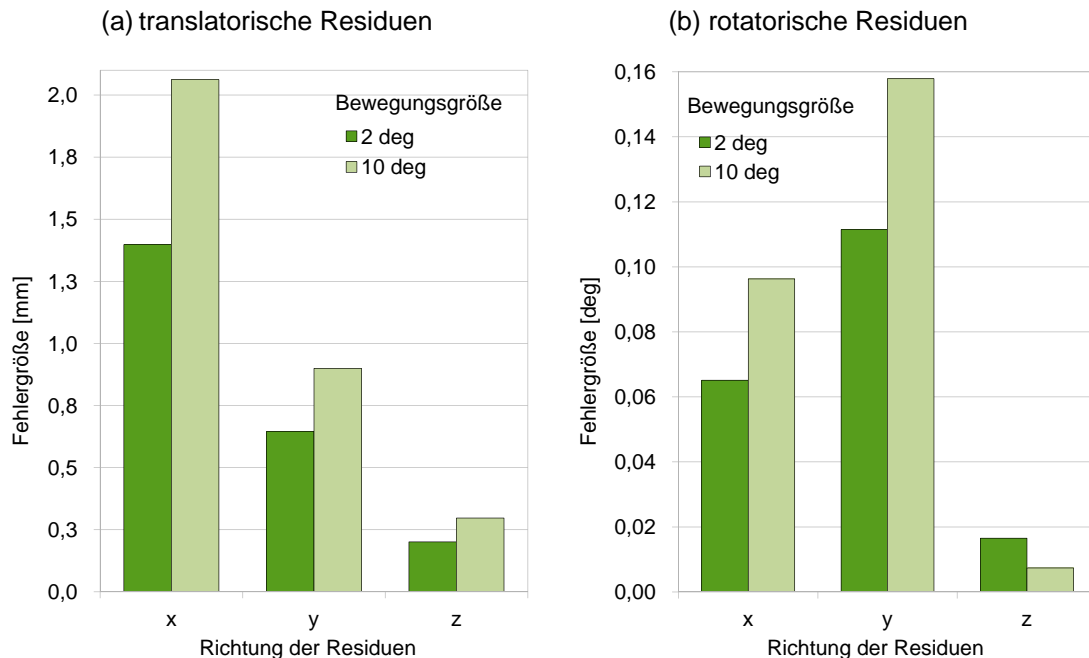


Abbildung 6.4 Rotation um y-Achse, Abhängigkeit Residuen von Bewegungsgröße

Vergleicht man die beiden Rotationsrichtungen miteinander, fällt auf, dass die Genauigkeiten der Rotationen um die y-Achse deutlich schlechter ausfallen. Das gilt für die translatorischen und rotatorischen Residuen. Eine Erklärung für den starken Anstieg der Residuen liegt in dem geometrischen Aufbau der texturierten Flächen. Die Korkfläche, auf die die Stereokameras gerichtet sind, ist zu ca. 2/3 parallel zur Bildebene der Kameras. Es werden also deutlich mehr korrespondierende Corners in der x-y-Ebene der Kameras detektiert. Der Algorithmus detektiert durch die planare Verteilung der Features Rotationen um die x- oder y-Achse schlechter, da die Bewegung mit einer Translation in die x- oder y-Richtung verwechselt werden kann. In Abbildung 6.4 ist zu erkennen, dass der größte translatorische Fehler in die x-Richtung geht, der größte rotatorische um die y-Achse. Eine kleine Verdrehung um die y-Achse hat im Bild etwa den gleichen Einfluss, wie eine translatorische Verschiebung in x-Richtung. Daher werden diese beiden Bewegungen leicht miteinander verwechselt und es entstehen Fehler in der Bewegungsschätzung. Je größer die Rotation wird, desto weiter steigen auch hier die Residuen an.

6.2 Abhängigkeit der Genauigkeit von der Anzahl der korrespondierenden Corners

Die Genauigkeit der Bewegungsschätzung hängt maßgeblich von den Eigenschaften der korrespondierenden Corners ab. Dieser Abschnitt soll dabei den Einfluss der Anzahl der korrespondierenden Corners bei der Schätzung aufzeigen. Die Anzahl kann durch Anpassung des Grenzwerts für die Feature-Detektion, wie in Abschnitt 5.2 beschrieben, beeinflusst werden. Im folgenden Abschnitt werden die Daten verwendet, deren Grundlage eine bildfüllende texturierte Fläche ist. Der Abstand der Stereokameras zu den Flächen liegt bei 650mm. Für eine bestmögliche Schätzung wird aus den bisherigen Erfahrungswerten von ca. 100 detektierten Features ausgegangen. Um Fehler zu provozieren, wird die Anzahl der korrespondierenden Corners schrittweise auf ca. 10 gesenkt.

6.2.1 Translatorische Bewegungen

Abbildung 6.5 visualisiert eine translatorische Bewegung von 10mm entlang der x-Achse. Die Genauigkeit der Bewegungsschätzung wird hier in Abhängigkeit der zur Bewegungsschätzung verwendeten Corners gezeigt. Am Beispiel dieser Translation ist zu erkennen, wie mit abnehmender Anzahl der für die Berechnung verwendeten Features die translatorische und rotatorische Genauigkeit der Schätzung schlechter wird. Die Datentabelle zeigt, dass diese Tendenz der Genauigkeit für alle Bewegungsgrößen in x- sowie in z-Richtung gilt (vgl. Anhang, A.2). Zwischen 30 und 100 Cornern verändert sich die Genauigkeit der Schätzung nicht signifikant. Ab weniger als 30 Cornern steigt die Größe der Residuen sprunghaft an. Das bestätigt die Annahme aus Abschnitt 4.4. Weniger als 10 korrespondierende Corners über das komplette Bild reichen für eine zuverlässige Bewegungsschätzung nicht mehr aus. Auffällig ist allerdings, dass bei den rotatorischen Fehlern ein systematischer Ausreißer in z-Richtung auftaucht. Das Verhalten wurde in Kapitel 6.1.1 bereits erläutert. Die Annahme eines systematischen Fehlers wird hier mit weiteren Daten fundiert. Wie aus der Datentabelle hervorgeht, gilt das auch für alle weiteren Bewegungen in x-Richtung mit variierender Cornerzahl (vgl. Anhang, A.2). Abbildung 6.6 entspricht den Daten von Abbildung 6.5(b).

Hier wurde zusätzlich die Standardabweichung der visuellen Odometrie zum Mittelwert der Bewegungsschätzung der geschätzten rotatorischen Bewegungen mit eingezeichnet. Die Standardabweichung der z-Werte ist für alle Datensätze

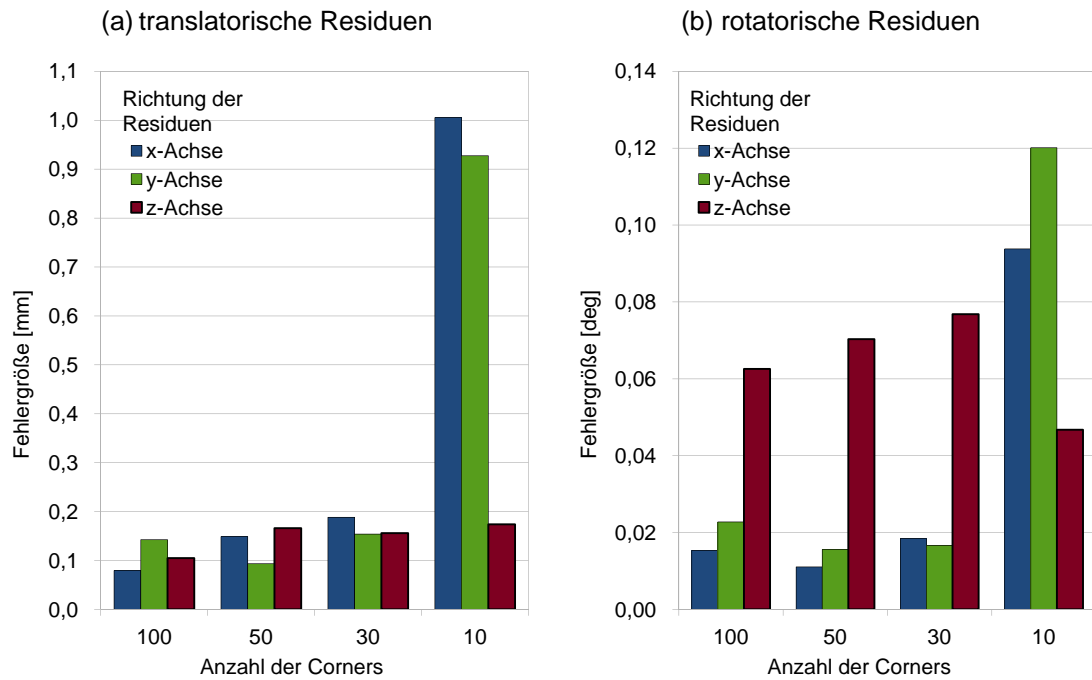


Abbildung 6.5 Translation in x-Richtung, Abhängigkeit Residuen von Corneranzahl

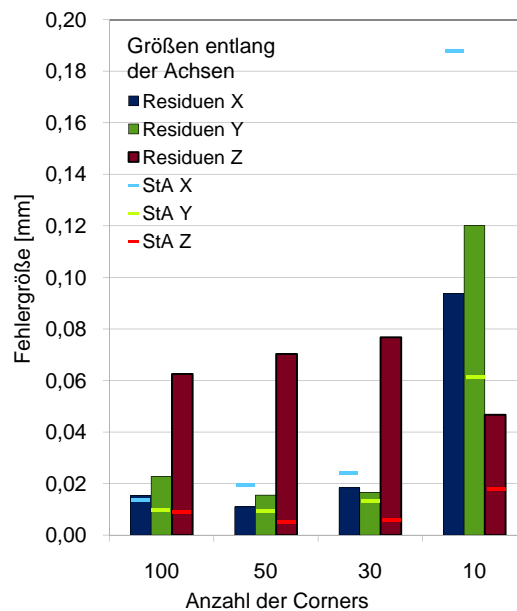


Abbildung 6.6 Translation in x-Richtung, Abhängigkeit Residuen von Corneranzahl mit Standardabweichung der visuellen Odometrie zu Mittelwert

durchgehend sehr klein. Daraus lässt sich schließen, dass die Bewegungsschätzung überall sehr eindeutig ist. In der weiteren Auswertung fällt dieser Fehler nicht signifikant auf. Deshalb kann die Auswertung auch ohne bekannte Ursache für dieses systematische Verhalten fortgeführt werden. Für die folgenden Abbildungen wird auf den Datensatz mit 10 Corners für die Berechnung verzichtet.

Abbildung 6.7 vergleicht die Residuen einer translatorischen Bewegung entlang der z-Achse. Dabei werden die Anzahl der verwendeten Corners, die Größe der Bewegung und die Verteilung der Fehler entlang der x-, y- und z-Achsen dargestellt. Sowohl bei

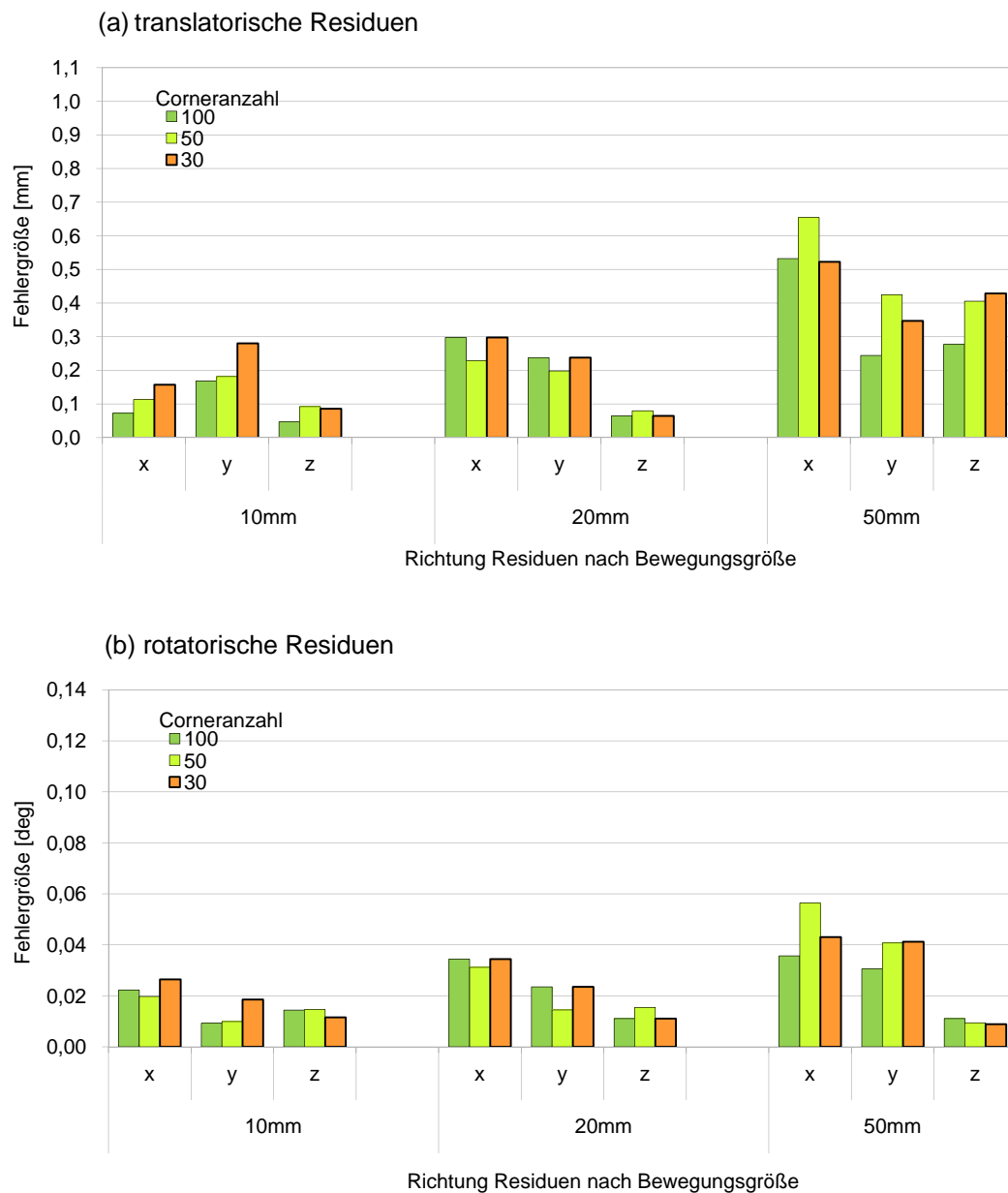


Abbildung 6.7 Translation in z-Richtung, Abhängigkeit Residuen von Corneranzahl, sortiert nach Richtung und Bewegungsgröße

den translatorischen, als auch bei den rotatorischen Residuen ist zu erkennen, dass sich die Genauigkeit der Schätzung mit 100 Cornern nur minimal von der mit 30 unterscheidet. Die Residuen zeigen das gleiche Verhalten wie bei einer Translation in x-Richtung, welches anhand der Abbildung 6.5 bereits erläutert wurde. Der systematische Fehler in den rotatorischen Residuen ist hier allerdings nicht gegeben.

6.2.2 Rotatorische Bewegungen

Abbildung 6.8 stellt wieder exemplarisch das Verhalten der Residuen mit abnehmender Anzahl der Corners dar. An der Rotation von 2° um die z-Achse sieht man, dass auch hier die translatorischen und rotatorischen Residuen ab weniger als 30 Cornern stark zunehmen. Die gleiche Tendenz gilt auch für Rotationen um die z-Achse von 10° , sowie für die Rotationen von 2° und 10° um die y-Achse, wie die Datentabelle zeigt. (vgl. Anhang, A.2).

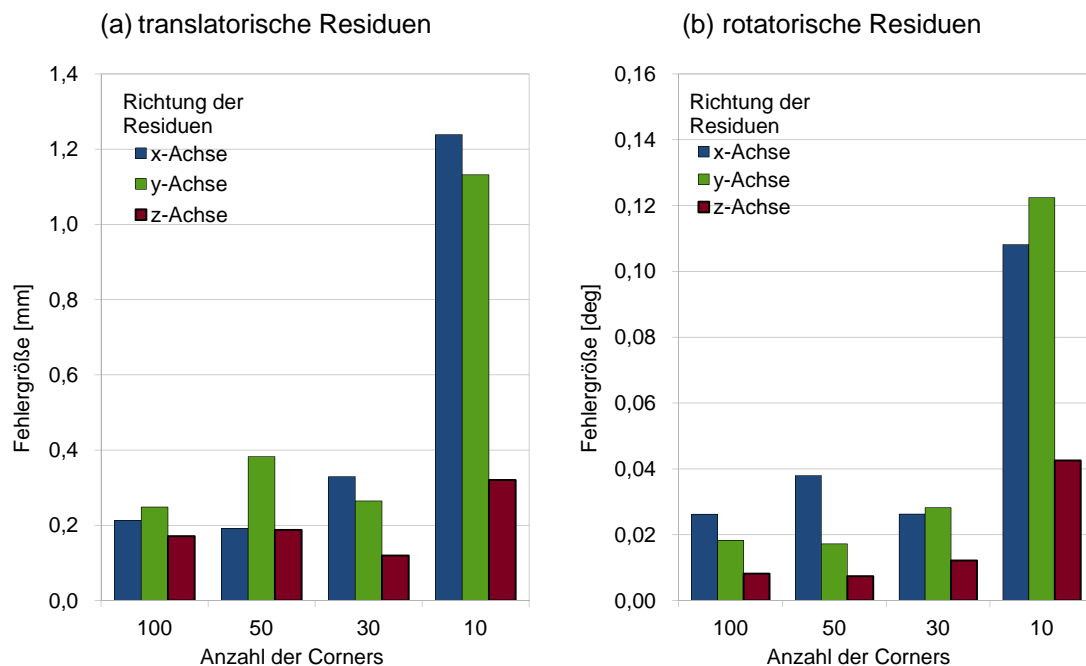


Abbildung 6.8 Rotation um z-Achse, Abhängigkeit Residuen von Corneranzahl

Abbildung 6.9 vergleicht die Residuen der rotatorischen Bewegungen um die z- Achse in Abhängigkeit der Anzahl der Corners für eine Bewegungsschätzung. Auch hier gilt, dass sich die translatorischen und rotatorischen Residuen zwischen 30 und 100 Cornern nicht signifikant ändern.

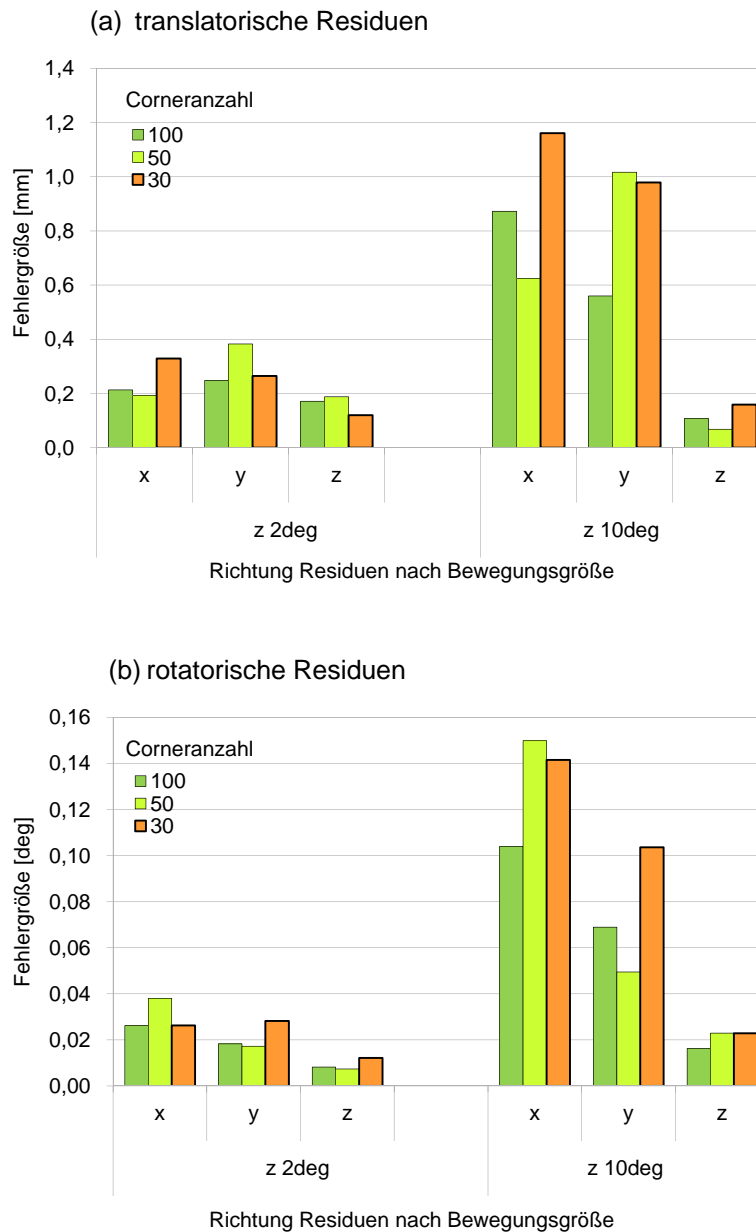


Abbildung 6.9 Rotation um z-Achse, Abhängigkeit Residuen von Corneranzahl, sortiert nach Richtung und Bewegungsgröße

6.3 Abhängigkeit der Genauigkeit von der Verteilung der korrespondierenden Corners im Bild

Eine weitere Eigenschaft der korrespondierenden Corners ist die Verteilung im aufgenommenen Bild der Szene. Die Verteilung wurde bei den Testreihen über die Größe der texturierten Fläche gesteuert. Die Abdeckung der Fläche in den Stereobildern beschreibt Tabelle 4.2. Dieser Abschnitt soll die Abhängigkeit der

Genauigkeit der Bewegungsschätzung des Algorithmus von der Verteilung der korrespondierenden Corners im Bild untersuchen. Durch die planare Form der texturierten Flächen von 30x30cm, 20x20cm und 8x8cm werden die Features nur in der x- und y-Ebene der Stereokameras gefunden. Für die Ausgangsdaten in den Abbildungen haben die Stereokameras einen Abstand von 650mm zu den texturierten Flächen und es werden für alle Bewegungsschätzungen in diesem Abschnitt ca. 100 korrespondierende Corners verwendet.

6.3.1 Translatorische Bewegungen

In Abbildung 6.10 wird die Abhängigkeit der Genauigkeit der Schätzung von der Verteilung der Features im Bild am Beispiel einer Translation entlang der x-Achse gezeigt. Es ist zu erkennen, dass die translatorische und rotatorische Genauigkeit sich

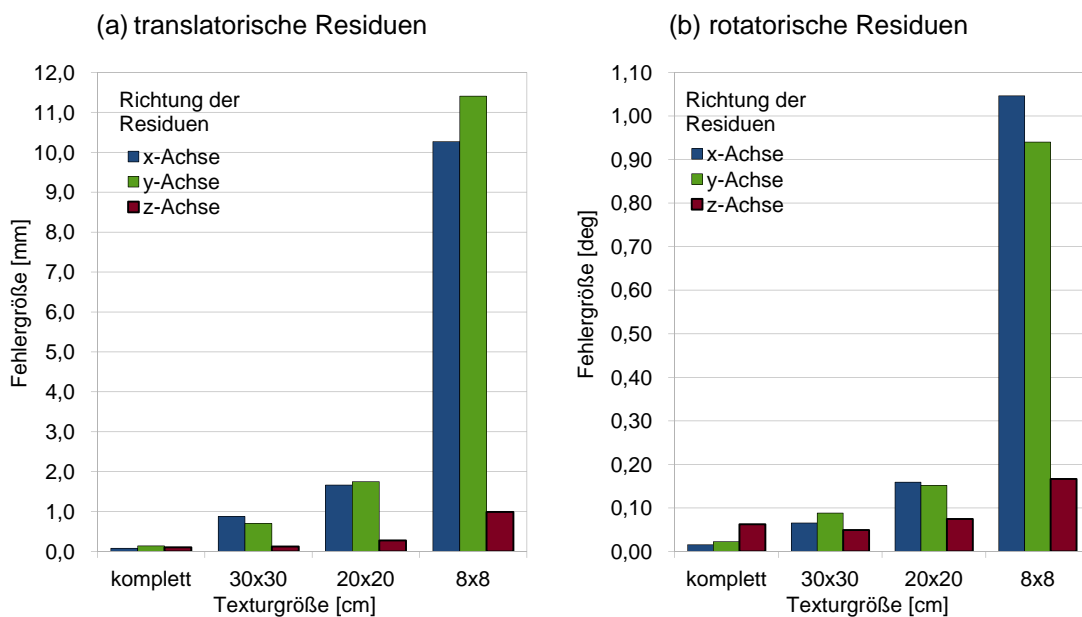


Abbildung 6.10 Translation in x-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche

mit der Verteilung der Corners deutlich ändert. Je weiter die Features über den Bildausschnitt verteilt sind, desto genauer wird die Bewegungsschätzung der Egomotion. Ab einer Größe der texturierten Fläche von 20x20cm werden die Ergebnisse der Egomotion sprunghaft schlechter. Deckt die Verteilung der Features also weniger als 10% der Bildfläche ab, kann keine zuverlässige Schätzung mehr geliefert werden. Der translatorische Fehler in x- und y-Richtung beträgt ca. 10mm. Dieses Verhalten zeigt sich in der Datentabelle auch für die restlichen translatorischen

Bewegungen (vgl. Anhang, A.1). Die folgenden Abbildungen stellen detaillierter die Abhängigkeit der Genauigkeit von der Verteilung der Features dar. Aufgrund der Übersichtlichkeit wurden die Daten der Bewegungen mit der Textur von 8x8cm nicht mit eingebunden, denn die Residuen sind hier deutlich größer, als bei den anderen Flächen.

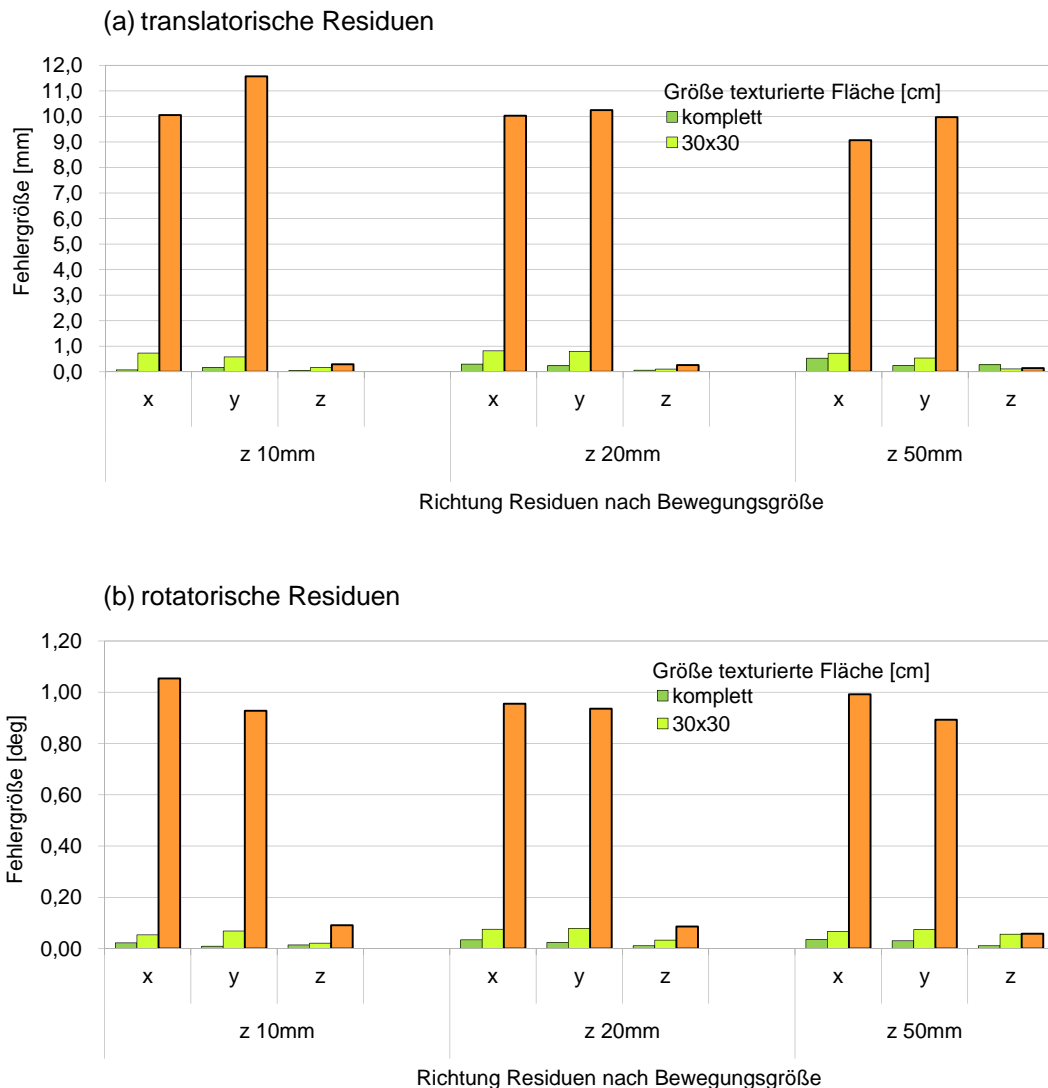


Abbildung 6.11 Translation in z-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche, sortiert nach Richtung und Bewegungsgröße

Abbildung 6.11 zeigt die translatorischen und rotatorischen Residuen einer Translation entlang der z-Achse in Abhängigkeit der Größe der Bewegung und der unterschiedlichen Texturen. Im Vergleich zu den anderen beiden Texturen steigen die translatorischen und rotatorischen Residuen in x- und y-Richtung bei der Textur

20x20cm sehr stark an. Der translatorische Fehler liegt bei ca. 10mm, der rotatorische bei ca. 1° unabhängig von der Bewegungsgröße. Eine zuverlässige Schätzung ist hier nicht mehr gegeben. Auch bei der Fläche von 30x30cm sind die Fehler in x- und y-Richtung am größten. Allerdings beträgt der translatorische Fehler hier nur ca. 0,8mm und garantiert somit noch eine gute Schätzung. Die großen Residuen in der x-y-Ebene entstehen dadurch, dass die aufgenommenen Flächen 30x30cm und 20x20cm planar sind. Die Features sind parallel zur Bildebene verteilt. Somit erscheint eine Rotation um die y-Achse im aufgenommenen Bild ähnlich wie eine Translation in x-Richtung und umgekehrt. Die Bewegungen werden verwechselt und erzeugen so diese Fehler. Je kleiner dabei die texturierte Fläche wird, desto stärker zeigt sich dieser Effekt.

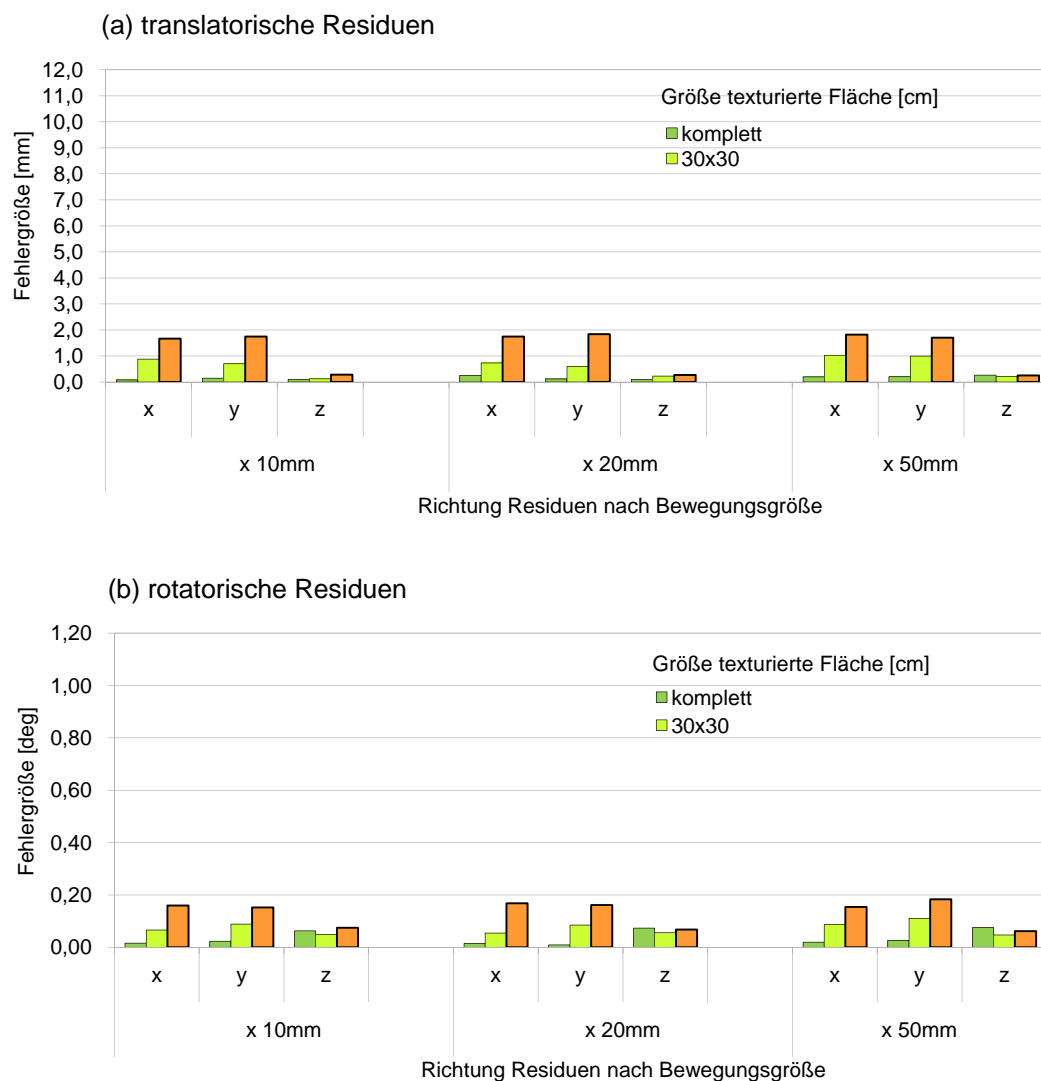


Abbildung 6.12 Translation in x-Richtung, Abhängigkeit Residuen von Größe der texturierten Fläche, sortiert nach Richtung und Bewegungsgröße

Die Residuen einer Translation in x-Richtung werden in Abbildung 6.12 beschrieben. Die Genauigkeit sinkt auch hier mit einer schlechteren Verteilung der Features im Bild. Das Gleiche gilt für die rotatorischen Fehler. Die translatorischen und rotatorischen Fehler der Texturen 30x30cm und 20x20cm sind in x- und y-Richtung wieder am größten, unabhängig der Größe der Bewegung. Die Residuen der Fläche 20x20cm betragen im Schnitt allerdings nur ca. 1,7mm. Dass die Residuen auch hier am größten in der x-y-Ebene sind, liegt wieder an den planaren Flächen der aufgenommenen Szenen.

Beim Vergleich der Residuen der beiden Translationen ist ein deutlicher Unterschied zu erkennen. Die Fehler bei einer Translation entlang der x-Achse sind bis zu Faktor sechs kleiner, als bei einer Translation auf das Objekt zu. Dabei spielt die Größe der Bewegung oder der texturierten Fläche keine Rolle. Der Grund für die größeren Residuen in Abbildung 6.11 wurde bereits in Abschnitt 6.1.1 genannt. Einerseits ist eine Translation entlang der optischen Achse generell schlechter zu bestimmen, als eine Bewegung parallel zur Bildebene. Andererseits zeigt sich das Skalierungsproblem. Bei einer Translation in z-Richtung findet eine Skalierung des Bildausschnitts statt, wobei das Suchfeld um die Corners bei dem Matching-Verfahren nicht angepasst wird.

6.3.2 Rotatorische Bewegungen

Für die rotatorischen Bewegungen wurden nur Daten mit der kompletten texturierten Fläche und der Fläche 30x30cm aufgenommen. Eine Konzentration der Features auf eine kleinere Fläche hätte zu schlechte Ergebnisse der Schätzung geliefert und wurde deshalb von vornherein nicht durchgeführt.

Abbildung 6.13 zeigt wieder exemplarisch an einer rotatorischen Bewegung von 2° um die z-Achse, wie sich die Residuen in x-, y- und z-Ebene verhalten. Die Genauigkeit der Schätzung nimmt mit kleiner werdender Fläche ab, wie in Abschnitt 4.4 erwartet. Anhand der Datentabelle ist zu erkennen, dass es die gleiche Tendenz auch bei allen anderen durchgeführten Rotationen gibt (vgl. Anhang, A.1).



Abbildung 6.13 Rotation um z-Achse, Abhängigkeit Residuen von Größe der texturierten Fläche

6.4 Abhängigkeit der Genauigkeit vom Abstand der Stereokameras zu der Szene

Dieser Abschnitt untersucht die Veränderung der Genauigkeit, wenn der Abstand der Stereokameras zu der Szene von 650mm auf 1050mm vergrößert wird. Grundlage für die folgenden Abbildungen sind die Daten mit ca. 100 Cornern für die Bewegungsschätzung, sowie eine komplett texturierte Fläche. Es wurde darauf geachtet, dass bei dem Vergleich der Daten nur die Testreihen gegenübergestellt werden, deren komplette Bildfläche texturiert war. Somit wird eine Änderung der Größe der texturierten Fläche verhindert.

6.4.1 Translatorische Bewegungen

Abbildung 6.14 zeigt die Residuen der Translationen und Rotationen bei den translatorischen Bewegungen in z-Richtung. Die Genauigkeitsuntersuchung bei einem Abstand von 650mm zu der texturierten Fläche wurde bereits in Abschnitt 6.1.1 durchgeführt. Die translatorischen und rotatorischen Residuen zeigen bei der Abstandsvergrößerung keine eindeutige Tendenz. Grundsätzlich überlagern sich hier zwei Effekte. Wie in Abschnitt 4.4 beschrieben wird mit einer Abstandsvergrößerung die translatorische Schätzung bei translatorischen Bewegungen schlechter. Betrachtet

man alle translatorischen Residuen bei einem Abstand von 1050mm, sind diese durchschnittlich etwas größer als bei 650mm. Auf der anderen Seite relativiert sich bei einer Abstandsvergrößerung das in Abschnitt 6.1.1 erörterte Skalierungsproblem bei translatorischen Bewegungen in z-Richtung. Bei größerem Abstand ändert sich bei einer Translation weniger im Suchbereich um die detektierten Features, der Algorithmus findet leichter korrespondierende Corners und eine Bewegungsschätzung wird besser. Das zeigt sich in Abbildung 6.14a. Bei einer Translation von 50mm und einem Abstand von 1050mm sind die Residuen im Verhältnis zu der 10mm-Translation ungefähr gleich geblieben. Bei dem Abstand von 650mm ist ein deutlicher Unterschied der Genauigkeiten zu erkennen. Des Weiteren ist bei größerem Abstand kein Anstieg mit ansteigender Bewegungsgröße zu erkennen. Die rotatorischen Residuen zeigen keine eindeutige Tendenz, wie in Abschnitt 4.4 erwartet.

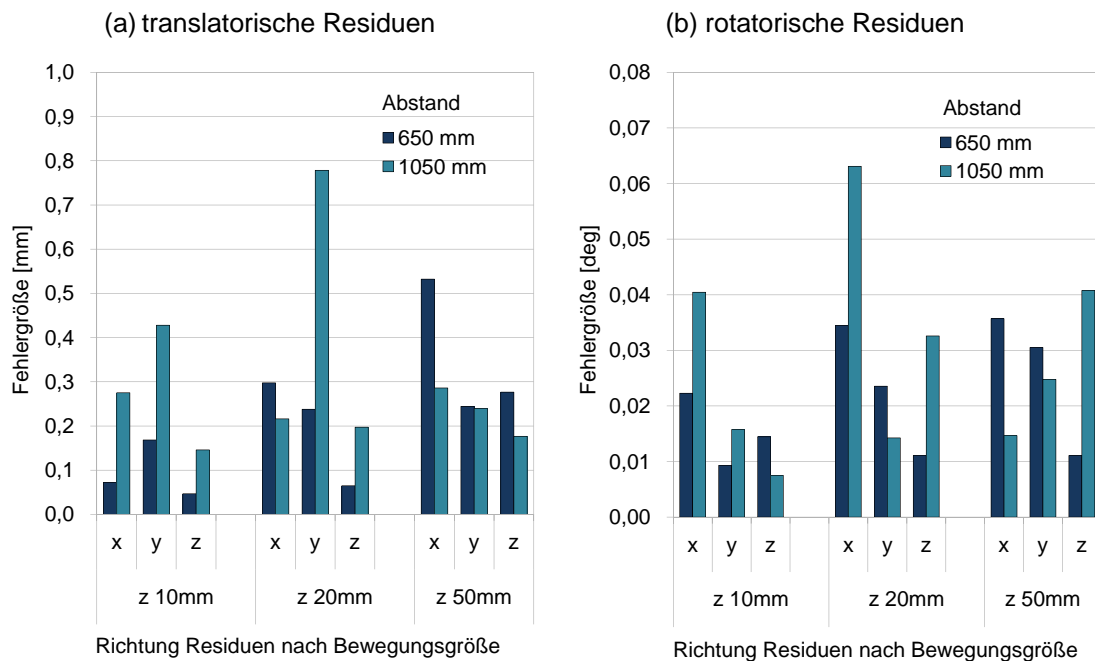


Abbildung 6.14 Translation in z-Richtung, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zur Szene, sortiert nach Richtung und Bewegungsgröße

Eine translatorische Bewegung in x-Richtung ist in Abbildung 6.15 dargestellt. Im Bild treten bei dieser Art von Bewegung keine Skalierungen auf, deswegen ändert sich im Verhältnis der Residuen der beiden Abstände weniger. In Abbildung 6.15a ist vor allem bei der Translation von 20mm ein Anstieg der Residuen bei einer Abstandsvergrößerung zu erkennen. Auch bei den anderen beiden Bewegungen ist ein

leichter Abfall der Genauigkeit zu sehen. Die rotatorischen Residuen sind beliebig verteilt und zeigen auch hier keine Tendenz.

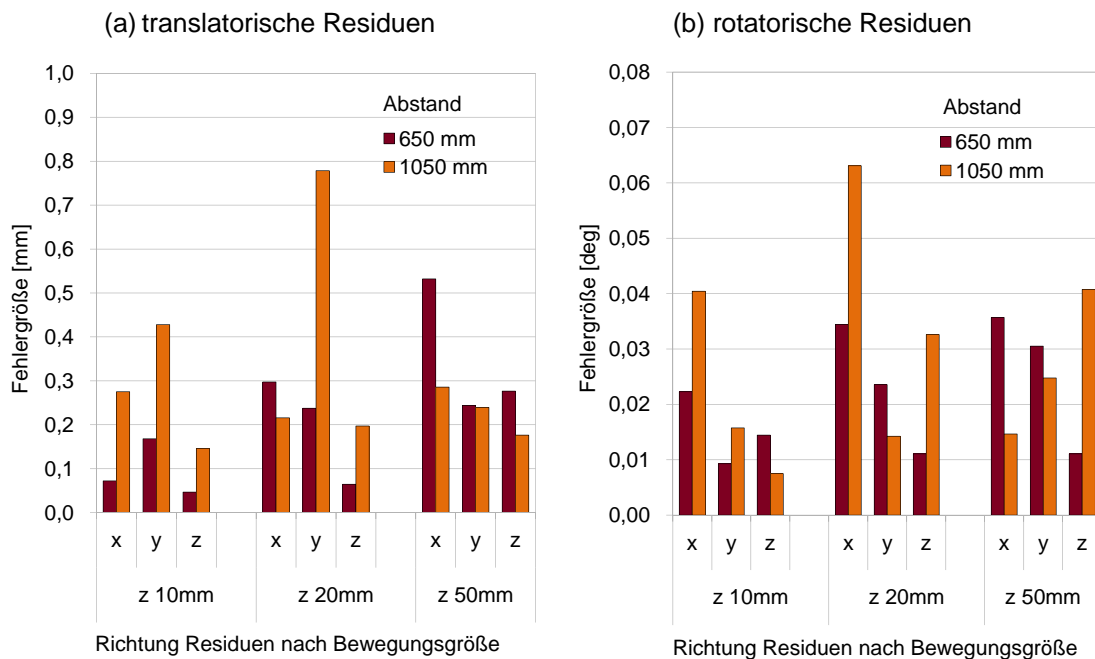


Abbildung 6.15 Translation in x-Richtung, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zur Szene, sortiert nach Richtung und Bewegungsgröße

6.4.2 Rotatorische Bewegungen

Die Abbildungen 6.16 und 6.17 visualisieren rotatorische Bewegungen der Stereokameras um die z- bzw. y-Achse in Abhängigkeit des Abstands. Abbildung 6.16b zeigt bei den Rotationen um 2° , dass sich der rotatorische Fehler wieder kaum ändert, wie in Abschnitt 4.4 erwartet. Bei einer Rotation um die y-Achse zeigt sich, dass die Residuen bei der Vergrößerung des Abstands deutlich besser werden. Wie in Abschnitt 4.4 erwähnt, sollte sich auch hier kein direkter Einfluss zeigen. Betrachtet man in Abbildung 6.17 die translatorischen Residuen entlang der x-Achse und die rotatorischen Residuen um die y-Achse bei einem Abstand von 650mm, sind das die größten Fehler. Dieses Verhalten wurde in Abschnitt 6.1.2 bereits erläutert. Die rotatorischen und translatorischen Bewegungen entlang dieser Achsen können leicht vertauscht werden, da die Texturfläche planar ist.

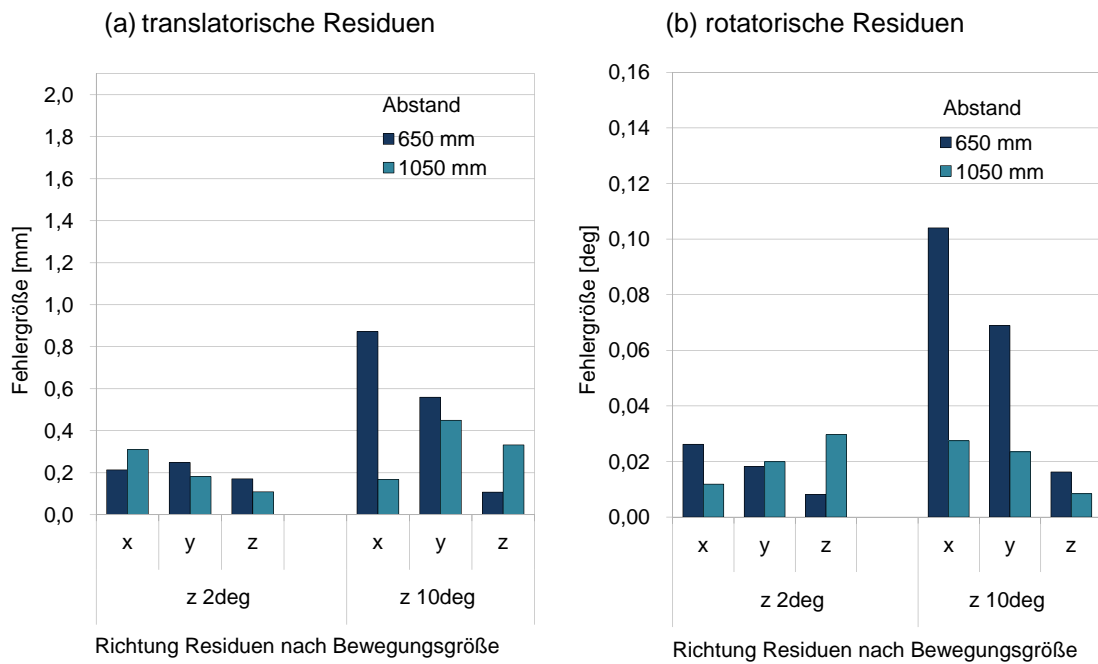


Abbildung 6.16 Rotation um z-Achse, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zu Szene, sortiert nach Richtung und Bewegungsgröße

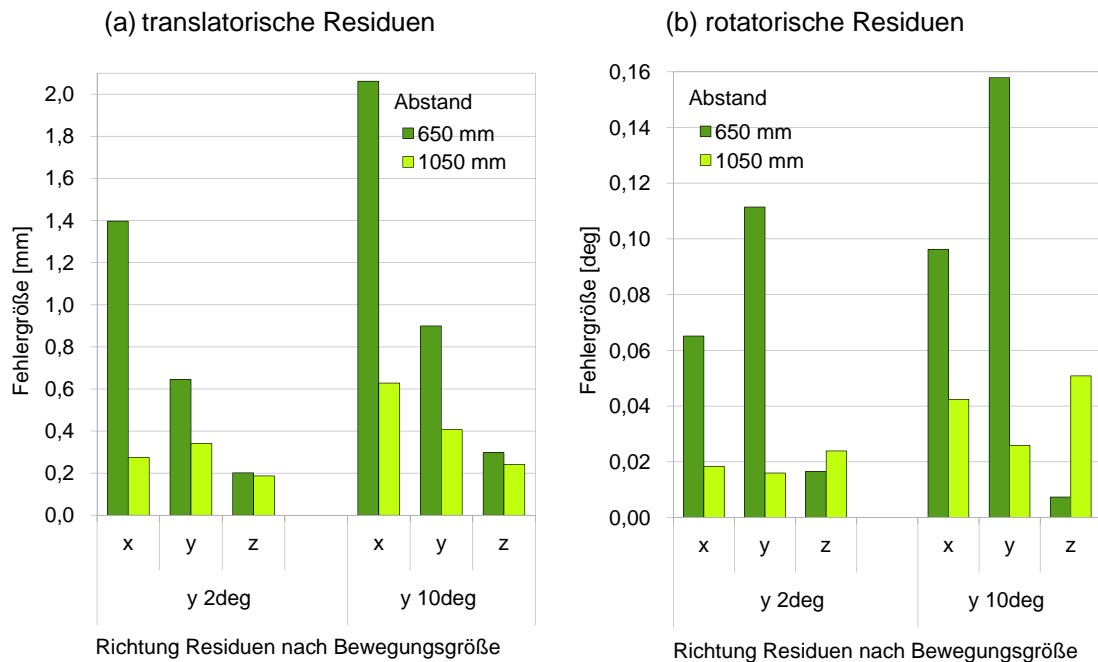


Abbildung 6.17 Rotation um y-Achse, Abhängigkeit Residuen von Vergrößerung des Abstands der Kameras zu Szene, sortiert nach Richtung und Bewegungsgröße

Abbildung 6.18 zeigt, dass durch das Design der Experimente bei einer Vergrößerung des Abstands auch mehr Dreidimensionalität der Szene in den Bildern vorhanden ist. Das überträgt sich auf die Genauigkeit der Schätzung, die dadurch erhöht wird. Der Zusammenhang zwischen Abstand der Kamera und Genauigkeit kann also nicht eindeutig gezeigt werden, weil andere Effekte die Genauigkeitsuntersuchung überlagern. Die translatorischen und rotatorischen Residuen entlang dieser Achsen sind bei einem Abstand von 1050mm daher deutlich kleiner. Das gilt für die Bewegungsgröße von 2 bzw. 10° .

(a) Abstand 650mm



(b) Abstand 1050mm



Abbildung 6.18 Vergleich der Zunahme der Dreidimensionalität bei Abstandsvergrößerung

6.5 Vergleich der tatsächlichen Fehler mit Fehlerschätzung des Egomotionalgorithmus

Die Funktionsweise der Fehlerschätzung wird in Abschnitt 2.4 erklärt. Die Schätzung der rotatorischen Fehler sind als Angle-Axis-Parameterfehler angegeben. Die geschätzten rotatorischen Fehler sind deswegen nicht direkt mit den berechneten Fehlern vergleichbar. Daher werden nur die translatorischen Fehler in diesem Abschnitt behandelt. Hier sollen die geschätzten Fehler des Algorithmus mit den tatsächlichen Residuen verglichen werden. Dabei soll herausgefunden werden, ob die Schätzung tatsächlich, wie in Abschnitt 4.4 angenommen, immer größer ist. Für die Abbildungen in diesem Abschnitt werden wieder die Daten mit maximaler Genauigkeit verwendet. Die Korktextur füllt den kompletten Bildausschnitt der Stereokameras, bei der Bewegungsberechnung werden ca. 100 Corners verwendet und der Abstand zu den Stereokameras beträgt 650mm.

6.5.1 Translatorische Bewegungen

Abbildung 6.19 vergleicht die tatsächlichen Fehler der translatorischen Bewegungen entlang der z-Achse und entlang der x-Achse mit der Schätzung des Fehlers der Egomotion. Dabei werden wieder die drei Translationsgrößen gegenübergestellt. In Abbildung 6.19(a) ist zu erkennen, dass die realen Residuen mit der Größe der Translation ansteigen. Die Schätzung zeigt diese Abhängigkeit nicht. Hier wird

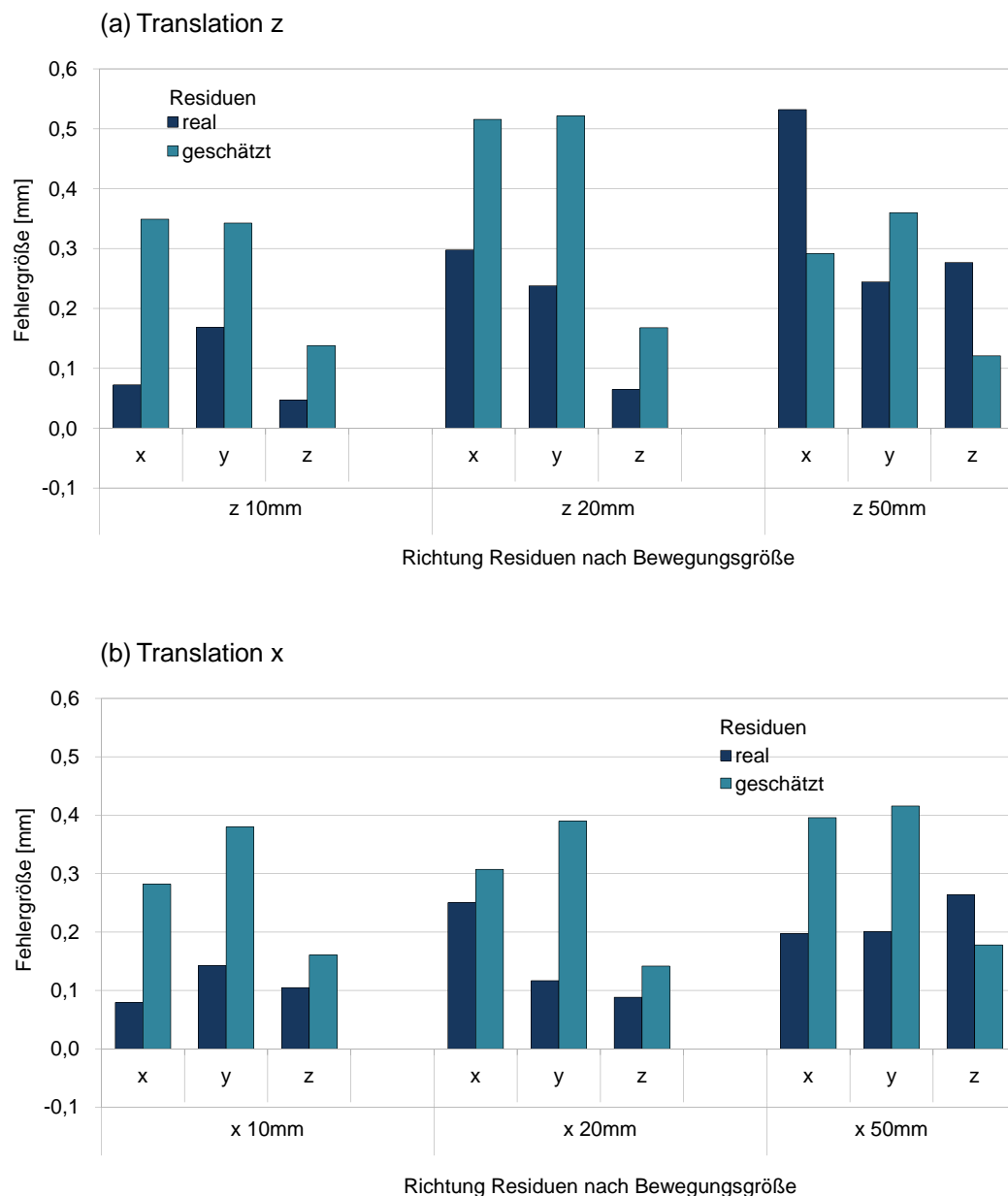


Abbildung 6.19 Translation in z- und x-Richtung, Vergleich von realen und geschätzten Residuen der visuellen Odometrie, sortiert nach Richtung und Bewegungsgröße

sichtbar, dass die Schätzung den Skalierungsfehler des Featurematching-Verfahrens aus Abschnitt 6.1.1 nicht mit einberechnet. Trotzdem sind die Fehlerschätzungen in den meisten Fällen größer als der tatsächliche Fehler. Das zeigt sich auch bei einer Translation in x-Richtung.

6.5.2 Rotatorische Bewegungen

Abbildung 6.20 zeigt den Vergleich der Residuen mit der Schätzung für die rotatorischen Bewegungen um die z-Achse und um die y-Achse. Auch hier zeigt sich in bei einer Rotation um die z-Achse, dass die Schätzung meistens schlechter ist, als der tatsächliche Fehler. Die Rotationsfehler beim Matching-Verfahren aus Abschnitt 6.1.2 werden von der Fehlerschätzung aber auch hier nicht mit einberechnet. Bei der Rotation um die y-Achse ist die Schätzung des Fehlers allerdings durchgehend besser. Auch hier ist eine Erklärung mit den bisher aufgenommenen Daten nicht möglich, da das Verhalten ähnlich wie in Abbildung 6.20a ist.

6.5.3 Zusammenfassung der Fehlerschätzung

Das Verfahren der Fehlerschätzung funktioniert in den meisten Fällen wie angenommen. Der geschätzte Fehler liegt über den tatsächlichen Residuen. Dass es bei manchen Fällen doch umgekehrt ist, kann verschiedene Gründe haben. Zum einen geht die Schätzung davon aus, dass keine fehlerhaften Korrespondenzen bei dem Featurematching-Verfahren detektiert werden. In der Realität treten aber dennoch ab und zu falsche Zuordnungen auf. Des Weiteren ist das Matching-Verfahren nicht skalierungs- und rotationsinvariant. Das wird bei der Schätzung auch nicht mit berücksichtigt.

Die Ausnahmen, die bei einer Rotation um die y-Achse auftreten, müssen in Zukunft noch weiter untersucht werden. Dafür ist allerdings eine neue Konzeption der Experimente notwendig. Der erste Schritt hier wäre eine Anpassung der texturierten Flächen, dass diese nicht mehr planar sind, sondern den Aspekt der Dreidimensionalität mit einbeziehen. Wie sich bei der Vergrößerung des Abstands gezeigt hat, werden die Schätzungen bei mehr dreidimensionaler texturierter Fläche besser.

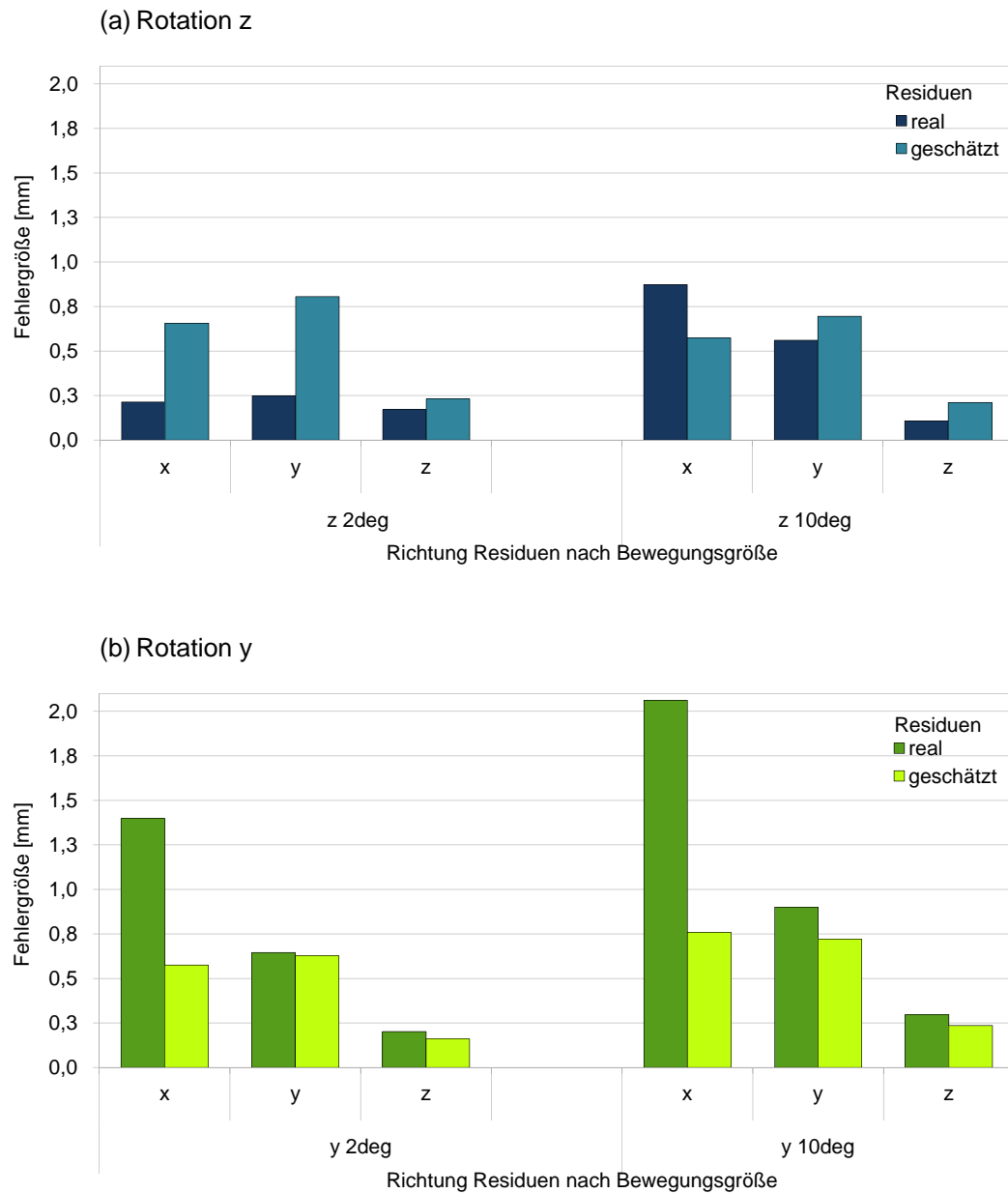


Abbildung 6.20 Rotation um z- und y-Achse, Vergleich von realen und geschätzten Residuen der visuellen Odometrie, sortiert nach Richtung und Bewegungsgröße

6.6 Zusammenfassung des Fehlerverhaltens des Egomotionalgorithmus

Die Auswertung der Experimente hat gezeigt, dass die Genauigkeiten der Bewegungsschätzungen des Egomotionalgorithmus von einigen Faktoren abhängen. Die visuelle Odometrie zeigt in einigen Fällen ein eindeutiges Fehlerverhalten. Dabei kommt es einerseits auf die Art der durchgeführten Bewegungen der Stereokameras

an, andererseits spielen die Eigenschaften der Features eine große Rolle für die Genauigkeit.

Bei allen Bewegungen ist der rotatorische Anteil um die z-Achse sehr genau bestimmbar. Das hat sich vor allem im Abschnitt 6.1.1 gezeigt. Führen die Stereokameras eine Bewegung parallel zu der aufgenommenen Szene durch, zeigt sich keine Abhängigkeit der Genauigkeit der Schätzung von der Bewegungsgröße. Wird eine translatorische Bewegung entlang der optischen Achse durchgeführt, werden die Schätzungen schlechter. Mit größerer Bewegung steigen auch die Residuen an. Dieser Effekt wird durch die nicht vorhandene Skalierungsinvarianz des Featurematching-Verfahrens ausgelöst.

In Abschnitt 6.1.2 hat sich herausgestellt, dass auch hier die Genauigkeit bei einer Rotation um die y-Achse von der Bewegungsgröße abhängt. Der Algorithmus verwechselt bei allen Bewegungen leicht translatorische Anteile in x-Richtung mit einer rotatorischen Bewegung um die y-Achse und umgekehrt, sowie translatorische Teile in y-Richtung mit rotatorischen um x. Je größer die Bewegung ist, desto stärker zeigt sich der Effekt. Diese Fehler werden durch die planare Gestaltung der texturierten Flächen deutlich verstärkt, wie sich in Abschnitt 6.4.2 gezeigt hat. Deswegen ist die Schätzung der Bewegung auch schlechter, als bei einer Rotation um die optische Achse. Bei einer rotatorischen Bewegung um die z-Achse steigen die Fehler mit der Größe der Bewegung. In diesem Fall liegt es aber an der nicht vorhandenen Rotationsinvarianz des Featurematching-Verfahrens.

In Abschnitt 6.2.1 und 6.2.2 hat sich gezeigt, dass die Genauigkeit der Bewegungsschätzung von der Anzahl der korrespondierenden Corners abhängt. Dabei steigen die Fehler zwischen 100 und 30 Corners für die Berechnung lediglich gering an. Werden aber nur noch 10 Corners oder weniger für die Schätzung verwendet, sind keine korrekten Ergebnisse mehr möglich. Das gilt für alle Bewegungsarten der Stereokameras, sowie für alle Bewegungsgrößen.

Die Experimente haben gezeigt, je geringer die korrespondierenden Features im Bild verteilt sind, desto schlechter wird die Schätzung. Die Aussagen über das Verhalten bei der Verteilung der Features wurde nur bei planaren Texturen überprüft. Bei einer Translation der Kameras parallel zu der Szene hat sich in Abschnitt 6.3.1 herausgestellt, dass eine Streuung der Features unter 10 % der Bildfläche keine vernünftigen Schätzungen mehr liefert. Bewegen sich die Kameras entlang der optischen Achse auf die Szene zu, werden die Schätzungen schon ab 20 % der Fläche deutlich schlechter. Das Gleiche gilt für rotatorische Bewegungen, wie in Abschnitt

6.3.2 erläutert. Diese Verhaltensweisen sind wieder alle unabhängig von der Größe der durchgeführten Bewegung.

Die Genauigkeiten der Bewegungsschätzung sind außerdem abhängig von dem Abstand der Stereokameras zu der aufgenommenen Szene. Die Schätzung der translatorischen Anteile der Bewegung bei einer Translation der Stereokameras wird schlechter, wie Abschnitt 6.4.1 gezeigt hat. Bei einer Translation entlang der optischen Achse relativieren sich die Skalierungen im Bild und die Schätzung wird bei einem größeren Abstand im Verhältnis besser. Da sich durch das Design der Experimente die dreidimensionale Komponente in der Szene geändert hat, ist es gerade bei den rotatorischen Bewegungen schwer zu beurteilen, ob die Fehler tatsächlich gleich geblieben sind.

Kapitel 7

Zusammenfassung und Ausblick

Das letzte Kapitel dieser Arbeit fasst die Ergebnisse in Abschnitt 7.1 zusammen gibt in Abschnitt 7.2 einen Ausblick für zukünftige Experimente und.

7.1 Zusammenfassung der Arbeit

In der vorliegenden Arbeit wurde der Egomotionalgorithmus von Dr. Heiko Hirschmüller auf seine Genauigkeit untersucht. Ziel dabei war es, die Abhängigkeit von Fehlern in der Bewegungsschätzung in Verbindung der detektierten Features im Stereokamerabild aufzuzeigen. Dafür wurde eine spezielle Testumgebung mit dem hochgenauen Roboterarm KR-16 von KUKA entwickelt und verschiedene Experimente durchgeführt. Getestet wurde die Abhängigkeit von der Bewegungsgröße der Kameras, von der Größe der texturierten Fläche im Bild, von der Anzahl der Corners für eine Bewegungsschätzung und der Abstand der Kameras zu der aufgenommenen Szene. Außerdem wurde die Fehlerschätzung des Algorithmus mit den tatsächlichen Fehlern verglichen, um diese zu verifizieren.

Die Ergebnisse haben gezeigt, dass der Algorithmus in vielen Situationen eine robuste Schätzung liefert, und so rechnet, wie bisher angenommen. Eine Zusammenfassung des Fehlerverhaltens zeigt, dass die Fehler wie erwartet durch die Verkleinerung der texturierten Flächen und die Verringerung der Anzahl der Corners entstehen. Außerdem ist die Genauigkeit auch abhängig von der Größe der durchgeführten Bewegung. Bei ansteigenden Translationen und Rotationen, vor allem entlang oder um die optische Achse steigen auch die Fehler. Das liegt an einer nicht vorhandenen Skalierungs- und Rotationsinvarianz bei dem Featurematching-Verfahren des Egomotionalgorithmus. Bei einer Vergrößerung des Abstands der Stereokameras zu der Szene steigen auch die Fehler an, da der Abstand der Features von der Kamera nicht mehr so genau rekonstruiert werden kann. Die Fehlerschätzung des Algorithmus hat sich bei den meisten Experimenten als richtig herausgestellt.

7.2 Ausblick

Die Experimente haben aber auch einige Fehlerverhalten des Egomotionalgorithmus aufgezeigt, die mit diesen Experimenten nicht erklärt werden konnten. Diese könnten mit weiteren Experimenten in Zukunft noch weiter untersucht werden.

Bei einer translatorischen Bewegung entlang der x-Achse hat sich in Abschnitt 6.1.1 und 6.2.1 ein systematischer rotatorischer Ausreißer um die z-Achse gezeigt, der mit dem bisherigen Wissenstand nicht erklärt werden kann. Um einen systematischen Fehler des Roboterarms zu überprüfen, müsste ein Experiment angefertigt werden, in dem zuerst eine Translation der Stereokameras entlang der x-Achse durchgeführt wird. Danach wird der Startpunkt des Roboterarms so verändert, dass alle Gelenke des Roboters in einer anderen Ausgangsstellung stehen. Somit könnte sicher gestellt werden, dass kein Gelenk des Roboters den gleichen Fehler macht wie bei der ersten Startposition. Von der neuen Startposition wird wieder eine Translation in x-Richtung mit der gleichen Bewegungsgröße wie zuvor durchgeführt. Tritt der Fehler in den Residuen dann nicht mehr auf, oder verändert sich, liegt es an dem Roboterarm des KR-16. Ist dies nicht der Fall, müsste der Fehler der visuellen Odometrie zugeordnet werden. Dann wären deutlich mehr redundante Bewegungsschätzungen notwendig, um eine genauere Aussage machen zu können.

Für die Experimente in dieser Arbeit wurde nur die planare Komponente der texturierten Flächen variiert. Bei einer Neugestaltung sollten die Flächen auch einen dreidimensionalen Anteil abdecken. Damit kann ausführlicher untersucht werden, wie sich die Verwechslung der Bewegungen entlang und um die x- und y-Achsen verhält. Der Einfluss der z-Richtung einer texturierten Fläche wurde nur in den Versuchen mit Vergrößerung des Abstands überprüft. Hier hat sich allerdings schon gezeigt, dass die Fehler an diesen Achsen zurück gehen, wenn die dreidimensionale Komponente der Fläche im Bild steigt.

Um das Verhalten bei einer Abstandsvergrößerung genauer untersuchen zu können, sollten zukünftige Experimente auch hier angepasst werden. Es hat sich gezeigt, dass eine Veränderung der 3D- Komponente der texturierten Fläche einen großen Einfluss auf die Genauigkeit der Bewegungsschätzung hat. Um nur den Effekt der Abstandsvergrößerung zu ermitteln, sollte sich die räumliche Eigenschaft der Fläche nicht verändern.

Ein weiterer Aspekt, der nach dieser Arbeit noch weiter analysiert werden könnte, ist das genaue Verhalten der Genauigkeit bei weniger als 30 korrespondierenden Corners für die Berechnung. Abschnitt 6.2 hat gezeigt, dass zwischen 10 und 30 Corners die

Fehler in der Schätzung sprunghaft ansteigen. Eine feinere Abstufung zwischen diesen Werten könnte zeigen, ob es einen Schwellwert gibt, ab dem die Ergebnisse deutlich schlechter werden, oder ob die Genauigkeit der Werte linear abnimmt. Dies ist für eine Performancesteigerung des Egomotionalgorithmus von Bedeutung. Je weniger korrespondierende Corners für die Berechnung verwendet werden, desto schneller kann eine visuelle Odometrie ausgegeben werden. Hier könnte noch ein besseres Gleichgewicht zwischen Rechengeschwindigkeit und Genauigkeit gefunden werden.

Bei der Verteilung der Features im Bild wurde in Abschnitt 6.3 auch ein sprunghafter Anstieg der Fehler detektiert. Hier könnte man das Verhalten der Genauigkeit zwischen 1% und 10% der Streuung im Bild noch genauer untersuchen. Dafür wären Experimente mit feineren Abstufungen der texturierten Flächen zwischen diesen Größen notwendig. Auch hier wäre es interessant, ob die Residuen ab einer bestimmten Verteilung sprunghaft ansteigen, oder linear mit der Verringerung der Fläche.

Literaturverzeichnis

- [1] CHILIAN, Annett: Stereokamerabasierte Navigation eines Krabbelroboters auf unebenem Gelände. Technische Universität Ilmenau, Fakultät für Maschinenbau, Diplomarbeit, 2008.
 - [2] HIRSCHMÜLLER, Heiko: Stereo Vision Based Mapping and Immediate Virtual Walkthroughs. Leicester, UK, De Montfort University, Diss., Juni 2003.
 - [3] HARRIS, C. ; STEPHENS, M.: A Combined Corner and Edge Detection. In: Proceedings of The Fourth Alvey Vision Conference, 1988, S. 147-151.
 - [4] CHILIAN, A.; HIRSCHMÜLLER,H.; GÖRNER,M : Multisensor Data Fusion for Robust Pose Estimation of a Six-Legged Walking Robot. In Proceedings of the International Conference on Intelligent Robotic Systems (IROS) 2011, 2011.
 - [5] KUKA Roboter GmbH: Spezifikation - Roboter KR 16 KS. URL: http://www.kuka-robotics.com/res/sps/e6c77545-9030-49b1-93f5-4d17c92173aa_Spez_KR_16_de.pdf. Abrufdatum 25.07.2011
 - [6] KUKA Roboter GmbH: KUKA System Software 5.5 - Bedien- und Programmieranleitung für Endanwender – Ver.: KSS 5.5 END V1 de, März 2008
 - [7] STROBL, K. H. ; HIRZINGER, G.: Optimal Hand-Eye Calibration. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006), Beijing, China, 2006
 - [8] DLR CalLab and CalDe - The DLR Camera Calibration Toolbox - Software available. URL: http://www.dlr.de/rm/desktopdefault.aspx/tabid-3925/6084_read-9201/. Abrufdatum: 25.07.2011
 - [9] Brief Tutorials DLR CalDe / DLR CalLab,
URL: http://www.dlr.de/rm-neu/desktopdefault.aspx/tabid-4853/6084_read-9196/.
Abrufdatum: 25.07.2011
-

Anhang

Erläuterung der Tabellenbeschriftung:

T_{res} translatorische Residuen [mm]

$\langle euler \rangle_{res}$ rotatorische Residuen [deg]

$\bar{T}_{err_{K1K2}}^{K1}$ geschätzte translatorische Fehler der visuellen Odometrie [mm]

$\sigma_{\bar{T}}$ Standardabweichung zum Mittelwert der Translationen der visuellen Odometrie [mm]

$\sigma_{\langle euler \rangle}$ Standardabweichung zum Mittelwert der Rotationen der visuellen Odometrie [deg]

σ_T Standardabweichung zum translatorischen Soll-Wert des KR-16 [mm]

$\sigma_{\langle euler \rangle}$ Standardabweichung zum rotatorischen Soll-Wert des KR-16 [deg]

\bar{d} mittlerer Abstand der Corners im Bild zum Schwerpunkt [mm]

mittlere Anzahl der korrespondierenden Corners für die Bewegungsschätzung

A.1 Daten mit variierender texturierter Fläche

Translatorische Bewegungen mit variierender texturierter Fläche

ca. 100 Corners für Bewegungsschätzung

Abstand Kameras - Szene 650mm

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
komplett																							
z 10mm	106	197	0,07	0,17	0,05	0,09	0,21	0,06	0,09	0,21	0,05	0,35	0,34	0,14	0,02	0,01	0,01	0,03	0,01	0,02	0,03	0,01	0,01
z 20mm	99	195	0,30	0,24	0,06	0,36	0,28	0,08	0,23	0,16	0,07	0,52	0,52	0,17	0,03	0,02	0,01	0,04	0,03	0,01	0,02	0,02	0,01
z 50mm	110	191	0,53	0,24	0,28	0,55	0,29	0,29	0,13	0,15	0,06	0,29	0,36	0,12	0,04	0,03	0,01	0,04	0,03	0,01	0,02	0,01	0,01
x 10mm	102	197	0,08	0,14	0,10	0,11	0,17	0,13	0,11	0,14	0,08	0,28	0,38	0,16	0,02	0,02	0,06	0,02	0,02	0,06	0,02	0,01	0,01
x 20mm	94	189	0,25	0,12	0,09	0,27	0,14	0,10	0,08	0,12	0,06	0,31	0,39	0,14	0,02	0,01	0,07	0,02	0,01	0,07	0,01	0,01	0,01
x 50mm	113	180	0,20	0,20	0,26	0,23	0,22	0,27	0,12	0,09	0,05	0,40	0,42	0,18	0,02	0,03	0,07	0,02	0,03	0,07	0,01	0,01	0,01
30x30 cm																							
z 10mm	104	129	0,73	0,58	0,16	0,90	0,74	0,20	0,85	0,73	0,14	2,19	2,11	0,39	0,05	0,07	0,02	0,07	0,08	0,03	0,07	0,08	0,02
z 20mm	95	127	0,82	0,80	0,11	1,00	1,00	0,14	0,91	0,76	0,11	2,17	2,04	0,39	0,08	0,08	0,03	0,09	0,10	0,04	0,07	0,08	0,02
z 50mm	91	130	0,72	0,54	0,11	0,87	0,68	0,14	0,57	0,48	0,12	1,92	1,85	0,35	0,07	0,07	0,06	0,08	0,09	0,06	0,05	0,06	0,02
x 10mm	103	131	0,88	0,70	0,12	1,06	0,93	0,15	1,00	0,93	0,11	2,32	2,24	0,44	0,07	0,09	0,05	0,08	0,10	0,05	0,09	0,09	0,02
x 20mm	96	131	0,73	0,60	0,22	0,90	0,77	0,27	0,82	0,77	0,16	2,37	2,20	0,45	0,05	0,08	0,06	0,07	0,10	0,06	0,07	0,07	0,01
x 50mm	96	129	1,02	1,00	0,21	1,20	1,24	0,24	0,82	1,04	0,13	2,32	2,32	0,50	0,09	0,11	0,05	0,11	0,13	0,05	0,09	0,07	0,02
20x20 cm																							
z 10mm	115	85	10,05	11,56	0,30	10,27	11,88	0,34	1,54	2,13	0,23	5,39	5,60	0,76	1,05	0,93	0,09	1,07	0,94	0,09	0,20	0,14	0,02
z 20mm	114	85	10,02	10,25	0,26	10,31	10,63	0,30	1,82	2,33	0,17	5,09	5,52	0,67	0,96	0,94	0,09	0,98	0,95	0,09	0,22	0,17	0,02
z 50mm	104	86	9,06	9,97	0,14	9,28	10,17	0,18	1,47	1,39	0,16	4,00	4,42	0,56	0,99	0,89	0,06	1,00	0,90	0,06	0,14	0,14	0,03
x 10mm	114	86	1,67	1,75	0,28	1,96	2,09	0,33	1,92	2,09	0,33	5,74	6,11	0,94	0,16	0,15	0,07	0,19	0,18	0,08	0,19	0,17	0,02
x 20mm	113	85	1,75	1,83	0,27	2,17	2,28	0,33	2,11	2,21	0,32	5,49	5,88	0,95	0,17	0,16	0,07	0,21	0,20	0,07	0,20	0,19	0,02
x 50mm	112	86	1,82	1,70	0,25	2,41	2,09	0,32	1,58	2,08	0,31	5,30	6,03	1,09	0,15	0,18	0,06	0,19	0,23	0,07	0,19	0,14	0,02
8x8 cm																							
z 10mm	102	33	7,29	9,57	10,68	8,56	11,59	10,87	7,79	9,05	1,05	17,24	21,03	2,67	0,89	0,66	0,07	1,07	0,76	0,09	0,85	0,72	0,08
z 20mm	97	33	7,89	7,75	10,18	9,60	10,42	10,40	8,34	9,93	1,49	14,50	18,07	2,59	0,73	0,72	0,06	0,98	0,87	0,08	0,95	0,79	0,08
z 50mm	99	33	12,97	10,87	10,01	14,54	12,70	10,21	6,80	7,92	1,24	13,45	17,85	2,68	1,03	1,24	0,09	1,20	1,38	0,11	0,79	0,67	0,09
x 10mm	92	33	10,27	11,41	0,99	12,99	13,78	1,29	12,27	13,77	1,22	22,06	25,85	3,10	1,05	0,94	0,17	1,25	1,17	0,19	1,26	1,11	0,13
x 20mm	98	34	9,56	12,29	1,66	13,41	15,91	2,04	12,98	14,06	1,50	21,07	21,99	3,11	1,13	0,87	0,14	1,45	1,21	0,17	1,29	1,18	0,12
x 50mm	93	32	8,02	18,22	2,31	9,99	21,11	2,73	8,49	12,21	1,74	24,50	21,34	3,98	1,69	0,74	0,13	1,93	0,90	0,17	1,12	0,77	0,13

Rotatorische Bewegungen mit variierender texturierter Fläche

ca. 100 Corners für Bewegungsschätzung

Abstand Kameras - Szene 650mm

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
komplett																							
z 2deg	105	204	0,21	0,25	0,17	0,26	0,31	0,19	0,24	0,30	0,09	0,66	0,81	0,23	0,03	0,02	0,01	0,03	0,02	0,01	0,03	0,02	0,01
z 10deg	101	201	0,87	0,56	0,11	0,91	0,65	0,13	0,21	0,36	0,08	0,57	0,69	0,21	0,10	0,07	0,02	0,11	0,07	0,02	0,03	0,02	0,01
y 2deg	109	204	1,40	0,65	0,20	1,44	0,71	0,22	0,29	0,30	0,08	0,57	0,63	0,16	0,07	0,11	0,02	0,07	0,11	0,02	0,03	0,03	0,01
y 10deg	104	170	2,06	0,90	0,30	2,13	0,99	0,32	0,43	0,40	0,12	0,76	0,72	0,24	0,10	0,16	0,01	0,10	0,16	0,01	0,04	0,04	0,01
30x30cm																							
z 2deg	111	115	0,28	0,60	0,11	0,33	0,68	0,13	0,31	0,37	0,07	2,42	3,16	0,50	0,05	0,03	0,00	0,06	0,03	0,01	0,03	0,03	0,01
z 10deg	111	117	0,36	0,97	0,31	0,44	1,07	0,32	0,34	0,52	0,06	2,17	2,63	0,45	0,05	0,03	0,01	0,06	0,04	0,01	0,05	0,03	0,01
y 2deg	97	116	0,75	0,36	0,17	0,85	0,47	0,19	0,39	0,36	0,08	1,95	1,70	0,38	0,04	0,06	0,04	0,05	0,06	0,04	0,03	0,04	0,01
y 10deg	102	117	2,23	0,61	0,38	2,32	0,80	0,39	0,54	0,63	0,08	2,19	2,14	0,34	0,07	0,18	0,01	0,08	0,19	0,02	0,06	0,05	0,01

A.2 Daten mit variierender Anzahl der Corners

Translatorische Bewegungen mit variierender Anzahl der Corners für Bewegungsschätzung
komplett texturierte Fläche
Abstand Kameras - Szene 650mm

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
ca. 50 Corners																							
z 10mm	48	203	0,11	0,18	0,09	0,14	0,23	0,12	0,12	0,22	0,11	0,51	0,53	0,18	0,02	0,01	0,01	0,02	0,01	0,02	0,02	0,01	0,01
z 20mm	46	201	0,23	0,20	0,08	0,26	0,24	0,10	0,15	0,13	0,07	0,40	0,40	0,14	0,03	0,01	0,02	0,03	0,02	0,02	0,02	0,02	0,01
z 50mm	50	199	0,65	0,42	0,41	0,68	0,47	0,42	0,19	0,23	0,08	0,48	0,63	0,17	0,06	0,04	0,01	0,06	0,04	0,01	0,03	0,02	0,01
x 10mm	51	197	0,15	0,09	0,17	0,19	0,13	0,18	0,14	0,11	0,07	0,37	0,45	0,21	0,01	0,02	0,07	0,01	0,02	0,07	0,01	0,02	0,01
x 20mm	57	190	0,35	0,16	0,12	0,37	0,19	0,14	0,11	0,18	0,07	0,45	0,51	0,17	0,02	0,01	0,07	0,02	0,01	0,07	0,02	0,01	0,01
x 50mm	55	183	0,22	0,22	0,40	0,24	0,24	0,41	0,09	0,09	0,07	0,51	0,50	0,23	0,02	0,02	0,07	0,02	0,02	0,07	0,01	0,01	0,01
ca. 30 Corners																							
z 10mm	35	200	0,16	0,28	0,09	0,19	0,35	0,10	0,17	0,29	0,10	0,65	0,69	0,21	0,03	0,02	0,01	0,03	0,02	0,01	0,03	0,02	0,01
z 20mm	32	195	0,30	0,24	0,06	0,36	0,28	0,08	0,23	0,16	0,07	0,52	0,52	0,17	0,03	0,02	0,01	0,04	0,03	0,01	0,02	0,02	0,01
z 50mm	28	196	0,52	0,35	0,43	0,60	0,43	0,45	0,36	0,34	0,14	0,67	0,80	0,24	0,04	0,04	0,01	0,05	0,05	0,01	0,04	0,04	0,01
x 10mm	36	192	0,19	0,15	0,16	0,25	0,19	0,18	0,19	0,18	0,10	0,53	0,72	0,26	0,02	0,02	0,08	0,02	0,02	0,08	0,02	0,02	0,01
x 20mm	30	184	0,24	0,20	0,21	0,30	0,23	0,24	0,25	0,20	0,11	0,73	0,89	0,26	0,02	0,02	0,06	0,02	0,03	0,06	0,02	0,03	0,01
x 50mm	32	181	0,44	0,28	0,41	0,47	0,30	0,42	0,16	0,07	0,06	0,83	0,61	0,30	0,03	0,01	0,07	0,03	0,02	0,07	0,01	0,02	0,01
ca. 10 Corners																							
z 10mm	9	162	0,47	1,61	0,30	0,72	2,20	0,53	0,72	2,15	0,49	3,27	5,93	1,45	0,18	0,05	0,01	0,24	0,08	0,02	0,23	0,08	0,02
z 20mm	11	176	0,57	2,40	0,18	0,88	3,59	0,28	0,86	2,64	0,26	2,90	5,41	0,74	0,27	0,05	0,02	0,39	0,09	0,03	0,29	0,09	0,02
z 50mm	8	174	0,92	1,05	0,45	1,08	1,30	0,47	0,64	1,12	0,13	4,74	8,66	1,24	0,13	0,09	0,02	0,15	0,10	0,02	0,13	0,07	0,02
x 10mm	11	177	1,01	0,93	0,17	1,15	1,23	0,21	0,58	0,94	0,18	2,56	4,96	0,96	0,09	0,12	0,05	0,12	0,13	0,05	0,10	0,06	0,01
x 20mm	12	169	0,37	0,45	0,24	0,46	0,52	0,28	0,43	0,38	0,16	1,60	1,83	0,50	0,05	0,05	0,05	0,06	0,06	0,05	0,04	0,04	0,01
x 50mm	11	150	0,44	0,27	0,46	0,55	0,33	0,50	0,46	0,21	0,17	3,20	2,12	0,88	0,02	0,06	0,05	0,03	0,08	0,06	0,02	0,05	0,01

Rotatorische Bewegungen mit variierender Anzahl der Corners für Bewegungsschätzung
komplett texturierte Fläche
Abstand Kameras - Szene 650mm

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
ca. 50 Corners																							
z 2deg	47	211	0,19	0,38	0,19	0,26	0,46	0,21	0,25	0,35	0,12	0,91	1,43	0,32	0,04	0,02	0,01	0,05	0,02	0,01	0,03	0,02	0,01
z 10deg	56	208	0,62	1,02	0,07	0,72	1,13	0,08	0,35	0,49	0,08	0,73	0,82	0,26	0,15	0,05	0,02	0,16	0,06	0,03	0,05	0,03	0,01
y 2deg	46	209	0,92	0,90	0,23	1,06	1,05	0,27	0,53	0,56	0,16	1,01	0,97	0,28	0,08	0,07	0,02	0,10	0,09	0,02	0,06	0,05	0,01
y 10deg	46	172	2,70	0,62	0,36	2,76	0,83	0,39	0,41	0,81	0,15	1,15	1,17	0,38	0,06	0,22	0,02	0,07	0,22	0,02	0,07	0,04	0,01
ca. 30 Corners																							
z 2deg	32	209	0,33	0,26	0,12	0,41	0,32	0,14	0,41	0,29	0,13	1,09	1,73	0,41	0,03	0,03	0,01	0,03	0,04	0,02	0,03	0,04	0,01
z 10deg	32	205	1,16	0,98	0,16	1,24	1,20	0,18	0,39	0,75	0,08	1,20	1,48	0,36	0,14	0,10	0,02	0,16	0,11	0,02	0,07	0,04	0,01
y 2deg	31	209	0,60	0,61	0,33	0,69	0,79	0,38	0,53	0,71	0,20	1,23	1,51	0,36	0,06	0,05	0,01	0,08	0,06	0,02	0,07	0,05	0,02
y 10deg	25	169	2,80	0,72	0,46	2,89	0,90	0,49	0,60	0,85	0,16	1,59	2,20	0,49	0,07	0,23	0,02	0,09	0,24	0,02	0,08	0,05	0,02
ca. 10 Corners																							
z 2deg	9	171	1,24	1,13	0,32	1,41	1,38	0,41	1,09	1,35	0,31	6,82	9,44	1,48	0,11	0,12	0,04	0,13	0,14	0,05	0,13	0,11	0,03
z 10deg	11	155	1,10	2,17	0,45	1,39	3,42	0,51	1,35	3,41	0,44	5,01	10,10	1,62	0,24	0,11	0,03	0,32	0,15	0,05	0,33	0,15	0,05
y 2deg	11	182	1,04	1,26	0,34	1,32	1,65	0,42	0,99	1,60	0,36	4,80	9,66	1,75	0,13	0,10	0,03	0,16	0,12	0,03	0,16	0,09	0,02
y 10deg	11	168	3,95	2,80	0,96	4,15	3,11	1,10	1,14	2,78	0,53	3,25	5,89	1,19	0,27	0,34	0,05	0,29	0,36	0,06	0,27	0,10	0,06

A.3 Daten mit Vergrößerung des Abstands der Stereokameras zu Szene

Translatorische Bewegungen Abstand Kameras - Szene 1050mm

komplett texturierte Fläche

ca. 100 Corners für Bewegungsschätzung

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
z 10mm	96	282	0,28	0,43	0,15	0,30	0,51	0,18	0,11	0,30	0,16	0,29	0,38	0,18	0,04	0,02	0,01	0,04	0,02	0,01	0,02	0,01	0,01
z 20mm	92	280	0,22	0,78	0,20	0,28	0,85	0,22	0,26	0,32	0,08	0,26	0,34	0,14	0,06	0,01	0,03	0,07	0,02	0,04	0,03	0,02	0,01
z 50mm	93	285	0,29	0,24	0,18	0,34	0,29	0,21	0,27	0,20	0,14	0,30	0,31	0,17	0,01	0,02	0,04	0,02	0,03	0,04	0,02	0,02	0,01
x 10mm	96	295	0,16	0,19	0,10	0,20	0,24	0,12	0,19	0,21	0,10	0,43	0,44	0,25	0,01	0,01	0,04	0,01	0,01	0,04	0,01	0,01	0,01
x 20mm	92	300	0,43	0,96	0,19	0,47	1,01	0,23	0,15	0,27	0,17	0,36	0,35	0,21	0,07	0,04	0,03	0,07	0,04	0,03	0,02	0,01	0,01
x 50mm	118	309	0,17	0,26	0,52	0,21	0,32	0,58	0,18	0,29	0,23	0,36	0,35	0,22	0,01	0,02	0,04	0,02	0,02	0,04	0,02	0,01	0,01

Rotatorische Bewegungen Abstand Kameras - Szene 1050mm

komplett texturierte Fläche

ca. 100 Corners für Bewegungsschätzung

	Corners		T_{res}			σ_T			$\sigma_{\bar{T}}$			\bar{T}_{err}^{K1}			$\langle euler \rangle_{res}$			$\sigma_{\langle euler \rangle}$			$\sigma_{\langle euler \rangle}$		
	#	\bar{d}	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z	x	y	z
z 2deg	103	289	0,31	0,18	0,11	0,37	0,23	0,13	0,31	0,20	0,13	0,42	0,37	0,24	0,01	0,02	0,03	0,01	0,02	0,03	0,01	0,02	0,01
z 10deg	103	284	0,17	0,45	0,33	0,21	0,49	0,36	0,21	0,17	0,13	0,38	0,44	0,29	0,03	0,02	0,01	0,03	0,03	0,01	0,01	0,01	0,01
y 2deg	107	294	0,27	0,34	0,19	0,31	0,42	0,23	0,29	0,27	0,15	0,23	0,28	0,13	0,02	0,02	0,02	0,02	0,02	0,03	0,02	0,02	0,01
y 10deg	109	286	0,63	0,41	0,24	0,67	0,48	0,27	0,23	0,31	0,12	0,27	0,31	0,16	0,04	0,03	0,05	0,05	0,03	0,05	0,02	0,01	0,01

Eigenständigkeitserklärung

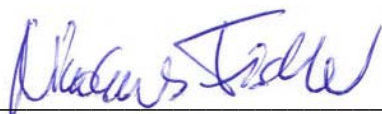
Nikolaus Fischer
15.09.1985
GI7 im SS 2011

München, den 26.07.2011

ERKLÄRUNG

gemäß § 35 Abs. 7 RaPO

Hiermit erkläre ich, dass ich die Bachelorarbeit selbständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benützt, sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.



(Unterschrift)